

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 5, 2004, 16:35:39 ; Search time 64 Seconds
(without alignments)
4489.694 Million cell updates/sec

Title: US-09-914-815A-10

Perfect score: 651

Sequence: 1 ggaagagagcgctggaac.....ccggccctactctgcc 651

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93.2	14.3	642	4	US-09-370-838-119 Sequence 119, App
2	69.6	10.7	352	4	US-09-702-705-1204 Sequence 1204, App
3	69.6	10.7	352	4	US-09-736-457-1204 Sequence 1204, App
4	68	10.4	207	4	US-09-702-705-998 Sequence 998, App
5	68	10.4	207	4	US-09-736-457-998 Sequence 998, App
6	64.6	9.9	907	4	US-09-620-312D-652 Sequence 652, App
7	45.6	7.0	1311	4	US-09-252-991A-1456 Sequence 1456, App
8	45.6	7.0	3396	4	US-09-252-991A-14676 Sequence 14676, App
9	45.6	7.0	3444	4	US-09-252-991A-15078 Sequence 15078, App
C 10	45.2	6.9	510	4	US-09-036-555B-163 Sequence 163, App
C 9	45.2	6.9	510	4	US-09-036-555B-163 Sequence 163, App
11	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
12	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
13	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
14	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
15	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
16	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
17	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
18	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
19	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
20	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
21	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
22	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
23	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
24	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
25	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
26	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App
27	43.2	6.6	745	1	US-08-469-569-163 Sequence 163, App

28	43.2	6.6	2003	1	US-08-469-526A-21 Sequence 21, App1
29	43.2	6.6	2003	2	US-08-734-591A-21 Sequence 21, App1
30	43.2	6.6	2003	2	US-08-469-660-21 Sequence 21, App1
31	43.2	6.6	2003	3	US-08-341-018-71 Sequence 71, App1
32	43.2	6.6	2003	3	US-08-470-335-21 Sequence 21, App1
33	43.2	6.6	2003	3	US-08-735-021-21 Sequence 21, App1
34	43.2	6.6	2003	3	US-08-734-664A-21 Sequence 21, App1
35	43.2	6.6	2003	3	US-08-470-339-21 Sequence 21, App1
36	43.2	6.6	2003	4	US-08-467-602-21 Sequence 21, App1
37	43.2	6.6	2003	5	PCT-US94-05083C-21 Sequence 21, App1
38	43.2	6.6	2003	5	PCT-US95-06846A-21 Sequence 21, App1
C 39	41.2	6.3	764	3	US-08-991-789A-283 Sequence 283, App
C 40	41.2	6.3	764	4	US-09-062-451-283 Sequence 283, App
C 41	41.2	6.3	764	4	US-09-289-198-283 Sequence 283, App
42	40.6	6.2	2846	3	US-09-613-182-5 Sequence 5, App1
43	39.8	6.1	2616	4	US-09-252-991A-16042 Sequence 16042, App1
C 44	39.8	6.1	2739	4	US-09-252-991A-16558 Sequence 16558, App1
C 45	39.4	6.1	615	4	US-09-501-115-21 Sequence 21, App1

ALIGNMENTS

```
RESULT 1
US-09-370-838-119
Sequence 119, Application US/09370838
Patent No. 644425
GENERAL INFORMATION:
APPLICANT: Lodes, Steven G.
APPLICANT: Mohamath, Roadoh
TITLE OR INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
TITLE OR INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.475C1
CURRENT APPLICATION NUMBER: US/09/370, 838
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/285,323
EARLIER FILING DATE: 1999-04-02
NUMBER OF SEQ ID NOS: 289
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 119
LENGTH: 642
TYPE: DNA
ORGANISM: Homo sapien
US-09-370-838-119

Query Match 14.3%; Score 93.2; DB 4; Length 642;
Best Local Similarity 66.3%; Pred. No. 6.8e-13;
Matches 134; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 125 GACCACTTATTCGTGCGGTTCACACCCCTTACTGTAATGTGTGCTCGGCTTC 184
DB 216 GACCACTTATTCGTGCGGTTCACACCCCTTACTGTAATGTGTGCTCGGCTTC 275
QY 185 CTGGCCCTGCTCTACTCCATCAAGCCGAGATCAGAGGTGTGTGTGCTGGAAGC 244
DB 276 ATGACATTCGCTCTACTCCATCAAGGTGTGTGTGCTGGAAGGTGTGCTGGAAGC 335
QY 245 GCCCGCGTTTGGCTTCAAGCCGAGATCAGAGGTGTGTGTGCTGGAAGGTGTGCTGGAAGC 304
DB 336 GCCCAGGCTTATCTCTCAAGCCGAGATCAGAGGTGTGTGTGCTGGAAGGTGTGCTGGAAGC 395
QY 305 GTGCGCGCTCTGCTGCTCGG 326
DB 396 CTATATCAATTCGCTCATCG 417

RESULT 2
US-09-702-705-1204
Sequence 1204, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
```

Query Match	10.7%;	Score 69.6;	DB 4;	Length 352
Best Local Similarity	56.6%;	Pred. No. 1.6e-07;		

RESULT 5
US-09-736-457-998
; Sequence 998, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong

Db 1055 CTGGGAGCGGTGTGATGGGGGATCTCTGTCTGACCATGTCACCTCTGTGTCTG 1114
Qy 365 GCCAAGACTCT 376
Db 1115 CCGGTGCTCTAT 1126

RESULT 8

US-09-252-991A-14676
Sequence 14676, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT FILING DATE: US/09/252,991A

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 14676

LENGTH: 3396

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-14676

Query Match 7.0%; Score 45.6; DB 4; Length 3396;

Best Local Similarity 48.8%; Pred. No. 0.077; Matches 123; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

Qy 125 GACCACTTATCTGTGCTGCTGTTGACGACCCCTTAACCTGTAATCTGTGCTGCTGCTTC 184
Db 3088 GCGCTGTGATGATCTCTTATCTCGCAACCTGTGAGAGACGACGCTGACACG 3147
Qy 185 CTGGCGCTGCTCTACTCATCAAGCCCGAATGCAAGAGTGTGTGACCTGGAACG 244
Db 3148 GCGGTGTGAGAGGGGCGCATCACCGCTGCGTGTGCTGATGACCGCTGTGTGCG 3207
Qy 245 GCCCGGCTTTTGGCTCCAAAGCCAAAGTCTACACATCTCGGCGGATGTGACGCTG 304
Db 3208 TCCCTCGTTTCGTGCGCATGAGCCCTGTGCGACCGGCGCGGAGGTGACGCGCG 3267
Qy 305 GTGCGGCACTGCTGCTCTCTGCGGCTGTGTGATGTGCTGACCTGACCTGCGGCTG 364
Db 3268 CTGGCGAGGTGTGATGCGGCGGATCTCTGTCTGACCATGACCTGTGTGTGCTG 3327
Qy 365 GCCAAGACTCT 376
Db 3328 CCGGTGCTCTAT 3339

RESULT 9

US-09-252-991A-15078/c

Sequence 15078, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT FILING DATE: US/09/252,991A

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 15078

LENGTH: 3444

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15078

Query Match 7.0%; Score 45.6; DB 4; Length 3444;

Best Local Similarity 48.8%; Pred. No. 0.077; Matches 123; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

Qy 125 GACCACTTATCTGTGCTGCTGTTGACGACCCCTTAACCTGTAATCTGTGCTGCTTC 184
Db 531 GCGCTGTGATGATCTCTTATCTCGCAACCTGTGAGAGACGACGCTGACACG 472
Qy 185 CTGGCGCTGCTCTACTCATCAAGCCCGAATGCAAGAGTGTGTGACCTGGAACG 244
Db 471 GCGGTGTGAGAGGGGCGCATCACCGCTGCGTGTGCTGATGACCGCTGTGTGCG 412
Qy 245 GCCCGGCTTTTGGCTCCAAAGCCAAAGTCTACACATCTCGGCGGATGTGACGCTG 304
Db 411 TCCCTCGTTTCGTGCGCATGAGCCCTGTGCGACCGGCGCGGAGGTGACGCGCG 352
Qy 305 GTGCGGCACTGCTGCTCTCTGCGGCTGTGTGATGTGCTGACCTGACCTGCGGCTG 364
Db 351 CTGGCGAGGTGTGATGCGGCGGATCTCTGTCTGACCATGCTGACCTGTGTGTCTG 292
Qy 365 GCCAAGACTCT 376
Db 291 CCGGTGCTCTAT 280

RESULT 10

US-09-252-991A-14818/c

Sequence 14818, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT FILING DATE: US/09/252,991A

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 14818

LENGTH: 510

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-14818

Query Match 6.9%; Score 45.2; DB 4; Length 510;

Best Local Similarity 49.2%; Pred. No. 0.068; Matches 119; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

Qy 125 GACCACTTATCTGTGCTGCTGTTGACGACCCCTTAACCTGTAATCTGTGCTGCTTC 184
Db 248 GCGCTGTGATGATCTCTTATCTCGCAACCTGTGAGAGACGACGCTGACACG 189
Qy 185 CTGGCGCTGCTCTACTCATCAAGCCCGAATGCAAGAGTGTGTGACCTGGAACG 244
Db 188 GCGGTGTGAGAGGGGCGCATCACCGCTGCGTGTGCTGATGACCGCTGTGTGCG 129
Qy 245 GCCCGGCTTTTGGCTCCAAAGCCAAAGTCTACACATCTCGGCGGATGTGACGCTG 304
Db 128 TCCCTCGTTTCGTGCGCATGAGCCCTGTGCGACCGGCGCGGAGGTGACGCGCG 69
Qy 305 GTGCGGCACTGCTGCTCTCTGCGGCTGTGTGATGTGCTGACCTGACCTGCGGCTG 364
Db 68 CTGGCGAGGTGTGATGCGGCGGATCTCTGTCTGACCATGCTGACCTGTGTGTCTG 9
Qy 365 GC 366
Db 8 CC 7

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 16:31:09 ; Search time 1971 Seconds

(Without alignments)
8027.506 Million cell updates/sec

Title: US-09-914-815A-10

Sequence: 1 ggaagagagcgctggaac.....cccgagccctaactctgcc 651

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estlin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_estc:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_estc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_ges_hum:*
- 18: em_ges_inv:*
- 19: em_ges_pln:*
- 20: em_ges_vrt:*
- 21: em_ges_fun:*
- 22: em_ges_mam:*
- 23: em_ges_mus:*
- 24: em_ges_pro:*
- 25: em_ges_rod:*
- 26: em_ges_pbg:*
- 27: em_ges_vrl:*
- 28: gb_gse1:*
- 29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	407	62.5	467	9	AA463818
2	345.4	53.1	568	9	AV589535
3	339	52.1	557	9	AV656559
4	334.6	51.4	541	12	BI849231

Result No.	Score	Query Match	Length	DB ID	Description
5	329.6	50.6	434	12	BM289220
6	295.4	45.4	720	11	BQ207465
7	286	43.9	739	11	AK003363
8	284.4	43.7	743	14	CB590125
9	283.2	43.5	485	9	AA033499
10	283	43.5	683	14	BY703620
11	280	43.0	725	13	BQ192887
12	279.2	42.9	728	14	CB321093
13	269.2	41.4	946	14	CB590001
14	261.6	40.2	443	9	AA000766
15	261.6	40.2	462	14	W75342
16	260.2	40.0	343	9	AW315021
17	260.2	40.0	449	9	AI322564
18	250.2	38.4	657	14	CB056845
19	246.6	37.9	339	14	W98365
20	245.8	37.8	666	13	BQ206122
21	224.6	34.5	414	14	W08017
22	213.2	32.7	371	13	BY337976
23	195	30.0	348	13	BY338389
24	193.4	29.7	281	13	BQ574249
25	191.8	29.5	505	9	AA274230
26	191.2	29.4	341	13	BY338388
27	183.8	28.2	304	13	BY336481
28	183.8	28.2	330	13	BY343164
29	183.8	28.2	330	13	BY343229
30	183.8	28.2	349	13	BY337306
31	183.8	28.2	367	13	BY337975
32	180.6	27.7	346	13	BY340550
33	174	26.7	270	9	AA460254
34	171	26.3	353	13	BY337972
35	167.8	25.8	451	9	AV589534
36	167.4	25.7	752	13	BU442736
37	163	25.0	439	10	BF651486
38	162.8	25.0	488	9	AW213665
39	161.4	24.8	321	14	W09955
40	161.4	24.8	498	10	BE688521
41	155	23.8	445	9	AV656558
42	153.6	23.6	299	13	BY336461
43	124.2	19.1	629	28	AZ361524
44	121.8	18.7	234	9	AA004139
45	121.2	18.6	291	9	AA016631

ALIGNMENTS

RESULT 1
LOCUS AA463818 467 bp mRNA linear EST 10-JUN-1997
DEFINITION z667d04.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
IMAGE:796519 5' similar to TR:G576631 G576631 14 KDA TRANSMEMBRANE
PROTEIN. mRNA sequence.
ACCESSION AA463818.1 GI:2188702
VERSION AA463818.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 467)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gaisel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, K., Stepec, M., Tan, F., Thelsting, B., White, Y., Wyllie,
'T., Waterston, R., and Wilson, R.
Washu-Merc EST Project 1997
TITLE Unpublished
JOURNAL
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LANT ; contact the


```

Db          520 ACCCCAGCCCC 531
|||||
RESULT 3
AV665659
LOCUS
DEFINITION
557 bp mRNA linear EST 28-NOV-2001
AV665659 Bos taurus brain fetus Bos taurus cDNA clone E1BR049H02
5', mRNA sequence.
ACCESSION
AV665659
VERSION
AV665659.1 GI:9924689
KEYWORDS
EST.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 557)
Takauega,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.
and Sugimoto,Y.
Establishment of a high throughput EST sequencing system using
poly(A) tail-removed cDNA libraries and determination of 36,000
bovine ESTs
Nucleic Acids Res. 29 (22), E108 (2001)
21570554
PUBMED
11713328
COMMENT
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazunugi@cocoa.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
FEATURES
source
1..557
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1BR049H02"
/issue_type="brain"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_1lb="Bos taurus brain fetus"
/notes="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"
BASE COUNT      89 a      214 c      158 g      91 t      5 others
ORIGIN
Query Match      52.1%; Score 339; DB 9; Length 557;
Best Local Similarity 83.9%; Pred. No. 3.8e-60;
Matches 407; Conservative 0; Mismatches 69; Indels 9; Gaps 2;
QY      CGCGCGTGAACCATGATGACGCGGATCCCGCGAGGACACCCGGGCCCCAGCGCCA 68
Db      12 CGAGGCTGGAACCATGATGACGCGGATCCCGCGAGGACACCCGGGCCCCAGCGCCC 71
QY      69 GCAAGGCGG-----GTGCCACACAGCCCTCACTGCGCGGCGCGACCCCGGCTC 122
Db      72 GCAAGGCGGACGCGGACGCGGCGGCTGAGCGTGGGGGGCTCGCGACCCCACTTC 131
QY      123 GAAACCACTGATCTGTGCGGTTCAGGACCTCTTACCTGAATCTGTGTGCTCGGCT 182
Db      132 GAAACCACTGATCTGTGCGGTTCAGGACCTCTTACCTGAATCTGTGTGCTCGGCT 191
QY      183 TCCTGGGCTGAGCTACTCATCAAGGCGCGAGATGGAAGTGTGAGTGAAGTGAAG 242
Db      192 TCCTGGGCTGAGCTACTCATCAAGGCGCGAGATGGAAGTGTGAGTGAAGTGAAG 251
QY      243 CGGCGGCGGCTTTTGCTCCAAAGCCAAAGTGTACCAATCTGCGCGGCGAGTGAAGCG 302
Db      252 CAGCGCGGCTCTCGGNTCCAAAGCCAAAGTGTACCAATCTGCGGCGAGATGGAAGCG 311

```

```

QY      303 TGGTCCCGCCACTGCTGCTCTGCGGCTGTGATGATGTCCTTGACACTGAGCCCGGC 362
Db      312 TGGTCCCGCCCGGCTCTGCTCTGCTGTGATGATGTCCTTGACACTGCTGCTGCGGC 371
QY      363 TGGCAAGGACTGTGCGCGGCTCTTCAAGACCAAGTTTATGACGGGGATATGACTGAC 422
Db      372 TGGCAAGGACTGTGCGCGGCTCTTCAAGACCAAGTTTATGACGGGGATATGACTGAC 431
QY      423 AGGCTGGGCTCTGATCTGAGGCACTAGCCCGACAGACTGACCGGCTGCTGCTCGG 482
Db      432 AGGCGGGGCTGTCTCTGCTG---TCCTGACAGGAGACCCCAACCCGAGGCTTACCCAG 488
QY      483 GGGCC 487
Db      489 CCCCC 493
RESULT 4
LOCUS
DEFINITION
541 bp mRNA linear EST 04-OCT-2001
B1849231
ACCESSION
B1849231
VERSION
B1849231.1 GI:15961750
KEYWORDS
EST.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 541)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrrenkrug,S.C., Bennett
,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Perteau,G., Holt,I., Karaymcheva,S., Liang,F., Quackenbush,J. and
Keeler,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
PUBMED
11282978
COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -mismatch 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCACGATCAGCAGC
Plate: 100 row: H column: 8
Seq primer: ATTAGTGTACACTATAG.
FEATURES
source
1..541
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/issue_type="pooled"
/lab_host="DH10B"
/clone_1lb="MARC 2BOV"
/notes="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, thymus,
semiteudonous muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
BASE COUNT      87 a      214 c      151 g      89 t
ORIGIN
Query Match      51.4%; Score 334.6; DB 12; Length 541;
Best Local Similarity 84.6%; Pred. No. 3e-59;
Matches 402; Conservative 0; Mismatches 64; Indels 9; Gaps 2;

```

QY 19 ACCATGACACGCGCTATCCCGGAGACACCCGCGCCCGCAGCAAGCG- 77
 DB 1 ACCATGACACGCTGTAACCCCGGAGACCCCGGCGCCCGCAGCGCGA 60
 QY 78 -----GTGCCACACAGCGCTCACTGGCGGCGCCGACCCCGCGTGAAGCACTT 132
 DB 61 CGGACACGCGCCACACGCGCTGACGCTGGGGGCTCCCGACCCCGCAGCAACTT 120
 QY 133 GATCTGGTGGTGTTCAGACACCTCTTACCTGAACTGTGTGCTCCGCTTCTGGGCT 192
 DB 121 GATCTGGTGGTGTTCAGACACCTCTTACCTGAACTGTGTGCTCCGCTTCTGGGCT 180
 QY 193 GGCTTACTCTCAAGGCGCGGATCAAGAGGTGTGTGACTGTGAAGCGCGCGCG 252
 DB 181 GGCTTACTCTCAAGGCGCGGATCAAGAGGTGTGTGACTGTGAAGCGCGCGCG 240
 QY 253 TTTTGGCTCCAAAGCCAAAGTGTCAACATCTGGCGCGGATGTGAGCGCTGTGGCG 312
 DB 241 TCTGGCTCCAAAGCCAAAGTGTCAACATCTGGCGCGGATGTGAGCGCTGTGGCG 300
 QY 313 ACTGCTGCTCTGGGCGCTGTGTGACTGTGCTGCTGCACTGGCGCGCGCGCAAGA 372
 DB 301 GGTCTGTCTGT 360
 QY 373 CTGTCGCGCTTCTTCAAGACCAAGTTGATGACGCGGACTGATGACAGGCTGGCT 432
 DB 361 CTGTCGCGCTTCTTCAAGACCAAGTTGATGACGCGGACTGATGACAGGCGGCG 420
 QY 433 CTGATCTGGGCGCTATGAGCGCGGACCACTGACCCCGCGCTGCTGCGGCGCG 487
 DB 421 CTGTCGCGCTTCTTCAAGACCAAGTTGATGACGCGGACTGATGACAGGCGGCG 472

RESULT 5 434 bp mRNA linear EST 28-DEC-2001
 LOCUS BM289220
 DEFINITION 530774 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION BM289220.1 GI:17998246
 VERSION
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
 Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.B., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-Mckown, C.G., Perle, G., Holt, I., Karaycheva, S., Liang, F., Quackenbush, J., and Keefe, J.W.
 Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)

TITLE
 JOURNAL MEDLINE
 PUBMED
 COMMENT
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smithth@mail.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.
 PCR primers
 FORWARD: AGGAACAGCTATGACCT
 BACKWARD: GTTTCAGCTACGACG
 Plate: 144 row: M column: 3
 Seq primer: ATTGATGACACTATAG.
 Location/Qualifiers
 1..434

BASE COUNT 66 a 166 c 126 g 76 t
 ORIGIN
 Query Match 50.6%; Score 329.6; DB 12; Length 434;
 Best Local Similarity 88.2%; Pred. No. 3.1e-58;
 Matches 372; Conservative 0; Mismatches 44; Indels 6; Gaps 1;
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_11b="MARC 3BOV"
 /note="Vector: PCMV SPORT6, Site 1: NotI, Site 2: SalI,
 library made from pooled tissue from marrow, alveolar
 macrophage, ovary, fetal semitendinosus muscle, and fetal
 longissimus muscle."
 20 CCATGACACGCGCTATCCCGGAGACACCCGCGCCCGCAGCAAGCGCG-- 77
 DB 1 CCATGACACGCTGTAACCCCGGAGACCCCGGCGCCCGCAGCGCGAAGCGCGAC 60
 QY 78 -----GTGCCACACAGCGCTCACTGGCGGCGCCGACCCCGCGTGAAGCACTTG 133
 DB 61 GGCAAGCGCGCCACACGCGCTGACGCTGGGCTCTCCGACACCCCGCAGCAACTTG 120
 QY 134 ATCTGCTGGTGTTCAGACACCTCTTACCTGAATCTGTGTGCTGCTGCTGCTGCTG 193
 DB 121 ATCTGCTGGTGTTCAGACACCTCTTACCTGAATCTGTGTGCTGCTGCTGCTGCTG 180
 QY 194 GCTTACTCTCAAGGCGCGGATCAAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 253
 DB 181 GCTTACTCTCAAGGCGCGGATCAAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
 QY 254 TTTGCTCCAAAGCCAAAGTGTCAACATCTGGCGCGGATGTGAGCGCTGTGCCGA 313
 DB 241 CTGCGCTCCAAAGCCAAAGTGTCAACATCTGGCGCGGATGTGAGCGCTGTGCCGA 300
 QY 314 CTGCTGCTCTGGGCGCTGTGTGACTGTGCTGCTGCACTGGCGCGCGCTGGCGCAAGAC 373
 DB 301 GTCCTGCTCTGGGCGCTGTGTGACTGTGCTGCTGCACTGGCGCGCGCTGGCGCAAGAC 360
 QY 374 TCTGCGCGCTTCTTCAAGACCAAGTTGATGACGCGGACTGATGACAGGCTGGCT 433
 DB 361 TCTGCGCGCTTCTTCAAGACCAAGTTGATGACGCGGACTGATGACAGGCGCGGCG 420
 QY 434 TG 435
 DB 421 TG 422

RESULT 6 720 bp mRNA linear EST 02-MAY-2002
 LOCUS BQ207465/c
 DEFINITION UT-R-DYI-cmo-e-14-0-UI-81 UT-R-DYI Rattus norvegicus cDNA clone
 ACCESSION BQ207465.1 GI:20423930
 VERSION
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 720)
 Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL MEDLINE
 PUBMED
 COMMENT
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized cartilage library cDNA library preparation. M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.reagen.com)

Seq primer: M13 Forward
SOLVA=yes.

MEDLINE	21085660	
PUBMED	11217851	
REFERENCE		
AUTHORS	5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II team.
TITLE		Analyses of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL		Nature 420, 563-573 (2002)
REFERENCE	6	(bases 1 to 739)
AUTHORS		Adachi,J., Alizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Caminici,P., Hakeda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotoh,K., Ishii,Y., Itoh,M., Iwata,M., Kaenuewa,T., Kato,H., Kawai,J., Kojima,Y., Komori,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishii,K., Nomura,K., Nunazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE		Direct Submission
JOURNAL		Submitted (10-Jul-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suenro-cho, Tsurumi-ku, Yokohama Kanagawa 230-0045, Japan [E-mail:genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216]
COMMENT		Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
		cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGCGCGCCGCAACTGAGCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with for full-length by sequence15' GAGAGAGAGAGATCCAGAGCTCAATTAATTAATTAACCCCCCCC 3'. cDNA was cleaved with XhoI and SctI. Cloning sites, 5' end: SctI; 3' end: XhoI. Host: SOLR.
FEATURES		Location/Qualifiers
SOURCE	1..739	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/strain="CS7BL/6J"	
	/db_xref="PANTOM:DB:1110003J06"	
	/db_xref="MGI:1906873"	
	/db_xref="taxon:10090"	
	/clone="111003J06"	
	/tissue_type="whole body"	
	/clone_lib="RIKEN full-length enriched mouse cDNA library"	
	/dev_stage="18-day embryo"	
misc_feature	62..466	
	/note="naemopoiesis related membrane protein 1 (MGDI MGI:1934923, GB AJ009781, evidence: BLASTN, 99%, match=469)	
	putative"	
	/db_xref="MGI:1934923"	
	721..726	
	/note="putative"	
polyA_site	739	
	/note="putative"	
BASE COUNT	160 a 228 c 186 g 165 t	
ORIGIN		
Query Match	43.9%	Score 286; DB 11; Length 739;
Best Local Similarity	68.1%	Pred. No. 3,5e-49;
Matches 436; Conservative	0; Mismatches 185; Indels 19; Gaps 2	
1	GGAGAGACGGCGCTGGAACCATGACACGGCGTATCCCGCGAGAGACACCGGAGCCCC 60	

[illegible]

```

source
1. 743
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30287396"
/tissue_type="embryonic limb, maxilla and mandible"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 136"
/Note="Vector: pCMV-Sport6.1.ccd; Site 1: EcoRV, Site 2:
NotI; Normalized, full-length enriched library from pool
of mouse embryonic limb, maxilla and mandible, embryonic
day 17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and
jaw equivalents from respective days). Cloned
directionally. oligo-dr primed
(5'-GACTGACTTAAATGCGGAGCGGCGCC(1715-3'. Size selected
for the 5kb fragments, average insert size 1.2 kb.
Normalization to Cot 7.5. Tissue contributed by David
Rowe; library constructed by Resgen, Invitrogen Corp.
Note: this is a NIH MGC library."

BASE COUNT      167 a      226 c      176 g      162 t      12 others
ORIGIN
Query Match      43.7%; Score 284.4; DB 14; Length 743;
Best Local Similarity 68.0%; Pred. No. 7.6e-49;
Matches 435; Conservative 0; Mismatches 186; Indels 19; Gaps 2;

QY 1 GGAAGAGACGGCCCTGGAACCCATGACACAGCGGTATCCCGGAGACACCCGGGCCCC 60
DB 16 GGAAGACACGGCTCTGGAACCCATGACATTCATATCCCGGAGACACCCGGGCTCC 75
QY 61 CAGCCACGACGAGCGCGT-----GCCACACAGCCCTCACTGGCGCGCCCGACCC 114
DB 76 ATATATCCGAGAGCGCTGATCTGACAGCCACACAGCCCTTCATGAGAAACCTTGCCCC 135
QY 115 CCGGCTGAGACCTTGAATCTGTGTGATTCAGACACCCCTTACCTGATCTGTGTTG 174
DB 136 TACACACAGATATCATGCTGTGTGTCTTCAAGACAGATGATCATGATCTGTGCTG 195
QY 175 CCGGCTGCTTCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 234
DB 196 CTTGGAATTCCTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 255
QY 235 CCGGAGAGCGCGCGGCTTGGCTCCAAAGCCAGTGTACATCTGCGCGCCGAT 294
DB 256 CTTGAGAGCTGACAGGACGATATGCTTCAAGCCAGTGTACATCTGCTGCTGCTGCTG 315
QY 295 GTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 354
DB 316 GTGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 375
QY 355 GGGCGGCTGCGGCAAGGACTGCGCGCTTCTTCAAGCAGGATTTGATGACGCGGACTA 414
DB 376 GTCAAGATTAGCCAAAGACTCTGCGCTTCTTCAAGCAGGATTTGATGAGAGACTA 435
QY 415 TGATGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
DB 436 TAACTAAGATTCGAGCTG-----TCCGTAACCGAGGAAACATGCTCA 482
QY 475 TGGCCCTGGGCGCAATATGACTCCCGGAGCTGCGCTCTTCTGTGCGGCTCTCAT 534
DB 483 GGTCACTGTGCGCAACACAGAGCTCTGCGGAGTTACAGCTTATAGACACCTCATGCC 542
QY 535 CCTGCGCCATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
DB 543 TGGCCAGCGGTGACCTCAAGAGCGGCGCTCTTGTCAAGATGTGTATTTGATGCTGCCC 602
QY 595 AGTCAGACCCCGGCGGTCTTCAACCTTAACCGAGAGTTCCC 634
DB 603 ACTCAATCTCTCAAGACTTAATTATTAACCTTAAGAGGCTCC 642

```

```

LOCUS      AA033499      485 bp      mRNA      linear      EST 22-AUG-1996
DEFINITION m144b03.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
clone IMAGE:466349 5' similar to SW:INIB RAT P26376
INTERFERON-INDUCIBLE PROTEIN. [1] ; mRNA sequence.
ACCESSION  AA033499
VERSION     AA033499.1
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 485)
REFERENCE   1 (bases 1 to 485)
AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterson,R.
TITLE      The WashU-HMI Mouse EST Project
JOURNAL    Unpublished
COMMENT    Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMGI Consortium (info@image.llnl.gov) for further information.
MGI:280165
Seq primer: -28M13 rev2 from Amerham
High quality sequence stop: 404.
FEATURES
source
1. 485
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:466349"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/clone_1ib="Soares mouse embryo NbME13.5 14.5"
/Note="Vector: pT73D-Pac (pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACCAATGAGATGGAGCGCGCGGAGAAATTTTTTTTTTTTTTTTTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos (total RNA provided by Minoru Ko, Wayne
State Univ., from 2 1; double-stranded cDNA was ligated to
Eco RI adaptor (pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bernaldo."

BASE COUNT      108 a      142 c      133 g      102 t
ORIGIN
Query Match      43.5%; Score 283.2; DB 9; Length 485;
Best Local Similarity 80.5%; Pred. No. 1.2e-48;
Matches 346; Conservative 0; Mismatches 78; Indels 6; Gaps 1;

QY 1 GGAAGAGACGGCGCTGGAACCCATGACACAGCGGTATCCCGGAGACACCCGGGCCCC 60
DB 51 GGAAGACACGGCGCTGGAACCCATGACATTCATATCCCGGAGAGACCCCGGCTCC 110
QY 61 CAGCCACGACGAGCGCGT-----GCCACACAGCCCTCACTGCGGCGCCCGACCC 114
DB 111 ATATATCCGAGAGCGCTGATGCTGACAGCCACAGCCCTCTTCAAGGAAACCTGCGCC 170
QY 115 CCGGCTGAGACCACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 174
DB 171 TACACACAGATATCATGCTGTGTCTTCAAGCAGATGTATCTGAATCTGTGCTG 230

```


of mouse embryonic limb, maxilla and mandible, embryonic day 17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and jaw equivalents from respective days). Cloned directionally, oligo-dt primed (5'-GACTTACTTCTAGATCGGAGGCGGCCCCC(17)15-3'. Size selected for the >1kb fragments, average insert size 1.2 kb. Normalization to Cot 7.5. Tissue contributed by David Rowe; library constructed by Resgen, Invitrogen Corp. Note: this is a NIH_MGC library."

BASE COUNT 181 a 219 c 173 g 157 t 8 others

Query Match 42.9%; Score 279.2; DB 14; Length 738;
Best Local Similarity 67.8%; Pred. No. 9e-48;
Matches 428; Conservative 0; Mismatches 184; Indels 19; Gaps 2;

10 GGGCTTGAACCCATGACACGCGGTATCCCGGAGAGACCCGGGCCCCACGCCAG 69
1 GGGCTTGAACCCATGACACCTTATCCCGGAGAGACCCCGGCGCTCATCATCCG 60
70 CAAGCCCGGT-----GCCACACAGCCCTCACACTGGCGGCCCCGACCCCGCTCG 123
61 CAAGCCCGGTATGACACCCACACAGCCCTTCCATGAGAACCTGCGCTACACACG 120
124 AGACCACTTATGATGT 183
121 AGATCATGT 180
184 CTTGGCGCTGGCTTATCTCATCAAGGCGGAGATCAAGAGTGTGTGTGTGTGTGTGT 243
181 CTTGGCGCTGGCTTATCTGTCAAGGCGGAGATCAAGAGTGTGTGTGTGTGTGTGT 240
244 GGGCGCGGT 303
241 TGCAGAGCAAGTATGCTTCAAGGCGGAGATCAAGAGTGTGTGTGTGTGTGTGTGT 300
304 GGTGCGCGCACTGT 363
301 GGTGCGCGCACTGT 360
364 GGGCAAGACCTGT 423
361 AGCCAAAGACTGT 420
424 GGT 483
421 GTTCCGAGCTG-----TCCCTGAACCGAGACACACATGTCAAGTCAAGTGT 467
484 GGGCCAAATCTGACTCCCGGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 543
468 TGCCCAACACCAAGCTCTGGAGGTTACAGCTTATAGCACTGTGTGTGTGTGTGTGT 527
544 ATCTGATCTGT 603
528 GTGACCTCAGAGGCGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 587
604 CCGGGGTCTTCAACCTTACCCGAGAGTTCCG 634
588 CTTCAAGACNTAATTATTAAGAGGCTCC 618

RESULT 13

CB590001

LOCUS

AGENCOURT 12601291 NIH MGC 136 Mus musculus cDNA clone

IMAGE:30287972 5', mRNA sequence.

ACCESSION CB590001.1 GI:29507857

VERSION EST.

KEYWORDS Mus musculus (house mouse)

ORGANISM Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 946)
AUTHORS NIH-MGC <http://img.mcg.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM324 row: k column: 21
High quality sequence stop: 657.
Location/Qualifiers

FEATURES

source

1..946
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30287972"
/tissue_type="embryonic limb, maxilla and mandible"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_136"
/note="Vector: pCMV-SPORT6.1.cdb; Site 1: Score; Site 2: Not; Normalized, full-length enriched library from pool of mouse embryonic limb, maxilla and mandible, embryonic day 17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and jaw equivalents from respective days). Cloned directionally, oligo-dt primed (5'-GACTTACTTCTAGATCGGAGGCGGCCCCC(17)15-3'. Size selected for the >1kb fragments, average insert size 1.2 kb. Normalization to Cot 7.5. Tissue contributed by David Rowe; library constructed by Resgen, Invitrogen Corp. Note: this is a NIH_MGC library."

BASE COUNT

221 a 275 c 233 g 217 t

Query Match 41.4%; Score 269.2; DB 14; Length 946;
Best Local Similarity 67.4%; Pred. No. 1.1e-45;
Matches 418; Conservative 0; Mismatches 183; Indels 19; Gaps 2;

21 CCATGACACAGCGGTATCCCGGAGAGACCCCGGCGCCCAAGCCCAAGAGCGGT- 79
1 CCATGACACCTTATATCCCGGAGAGACCCCGGCGCTCATCATCCGCAAGCTGATG 60
80 -----GCCACACAGCCCTTCACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 134
61 CTGACAGCCCAACAGCCCTTCCATGAGAACCTGCGCCCTTACACACAGATCATCATG 120
135 TGT 194
121 TGT 180
195 CTTATCTCATCAAGCGCGAGATCAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 254
181 TCACACTGTCTCAAGGCGGCGGAGACAGAAAGATGAGTGTGTGTGTGTGTGTGTGTGT 240
255 TTGGCTTCAAGCGAGT 314
241 ATGGCTTCAAGCGAGT 300
315 TGT 374
301 TGT 360
375 CTGCGCGCTTCTTCAAGACCAAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 434
361 CTGCGCGCTTCTTCAAGACCAAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
435 GATCTGGGGGACTAGGCGGAGCACTGACCCGAGGCTGTGTGTGTGTGTGTGTGTGTGT 494
421 G-----TCCCTGAACCGAGGACACATGTCAAGTCAAGTGTGTGTGTGTGTGTGT 467

OY	495	GATCTCCCGGAGCCTGGGCCCTTCCTGTGTGGGGGCCTCATCCCTGCCCCACCTTGATCT	554
Dd	468	AGCTCTGGGAGATTACAGACTTACTATAGCACTCTGCATCCCTGGCCAGCCTGTGACTCAGA	527
OY	555	GGGGCCCTCCAGCCCCAACATGGGACCCTTAAGGTGAACCAATCCAGTCCAGCCCAGGGCTTC	614
Dd	528	GGGGCCCTCTTTGTTCACGATGTGTATTGATNGCTGGCCCACTCAGATCCCTCAGACTTA	587
OY	615	ACCTTAACCCGAGAGTTCCC	634
Dd	588	ACTATAAACCCTAGAGGCTCC	607
<hr/>			
RESULT 14			
LOCUS	AA000766		
DEFINITION	AA000766	443 bp	mRNA linear EST 18-JUL-1996
ACCESSION	mg38a01.r1	Soares mouse embryo NM0E13.5 14.5 Mus musculus cDNA	
VERSIONS	INTERFERON-INDUCIBLE PROTEIN. [1], mRNA sequence.		
KEYWORDS	AA000766		
SOURCE	AA000766.1 GI:1436846		
ORGANISM	EST.		
REFERENCE	Mus musculus (house mouse)		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 443) Matra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterson,R.		
TITLE	The WashU-HMMI Mouse EST Project		
JOURNAL	Unpublished		
COMMENT	Contact: Marra M/Mouse EST Project WashU-HMMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel.: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGJ:260576 Seq primer: ETPrimer High quality sequence stop: 354. Location/Qualifiers		
FEATURES	source		
	1..443		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:426024"		
	/sex="unknown"		
	/tissue_type="embryo"		
	/dev_stage="13.5-14.5dpc total fetus"		
	/lab_host="DH10B"		
	/clone_id="Soares mouse embryo NM0E13.5 14.5"		
	/note="(vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15', TGTACCACTCTGAGTAGGAGGAGCGCGGAGAAATTTTTTTTTTTTTTTTTTTTTTTT T 3'), on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne Staele Univ., from 2 j]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3D vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Patima Bernaldo."		
BASE COUNT	96 a	135 c	115 g 97 t
ORIGIN			

Query Match	40.2%	Score 261.6	DB 9	Length 443
Best Local Similarity	78.5%	Pred. No. 3.4e-44		
Matches 328	Conservative 0	Mismatches 84	Indels 6	Gaps 1
QY	24	TGACACGCGGCTATCCCGCGAGGACACCCGCGGCCCCCGACCGCCAGCAGCAGCCGGT----	79	
DB	3	TGACACTTCAATATCCCGTGATGAGCCCGGGCTTCATCATCCCGCAAGGCTGATGCTG	62	
QY	80	--GCCCAACAGCCCTCACTGCGCGGCCCCCGCACCCCGGCTCTGAGACCACTTGATCT	137	
DB	63	CAGCCCAACAGCCCTCTCCATGGAACACCTGGCCCTTACACCAACGATCACTGCTCT	122	
QY	138	GGTCGGATTGAGCAACCTTACCTGGAATCTGTGTGTGCTCGGCTCTCTGGGCTGAGCT	197	
DB	123	GGTCGTCTTGAAGACATGTACTGAATCTGTCTGCTCTTGATCTCTGGGCTGAGCTC	182	
QY	198	ACTCCATCAAGCCCGAGATCAGAAAGGTGTGTGTGACCTTGAAACGCGCCGCGCTTTG	257	
DB	183	ACTGTGTCAAGGCCCGAGACAGAAAGATGTGCTGGAACTTGAGAGGCTCAAGGCAAGTATG	242	
QY	258	GCTCCAAAGCCAAATGCTTCAACAATCTCTGGCGCGGATGTGAGACGTGTGCTGCACTGC	317	
DB	243	GCTCCAAAGCCAAATGCTTCAACAATCTCTGGCTGCAATGTGACATGTGTGCCCCCATTCG	302	
QY	318	TGCTCTGGGCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG	377	
DB	303	TGCTCTGGGCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG	362	
QY	378	CCGCTCTTCTTCAAGCAAGATTGATGACGCGGACTATGACTGACAGGCTGGTCTTG	435	
DB	363	CGGCTCTTCTTCAAGCAAGATTGATGATGAGGAGACTATGACTGAGATTGCGAGCTTG	420	
RESULT 15				
W75342				
LOCUS				
DEFINITION	W75342	462 bp	mRNA	linear
ACCESSION	W75342			EST 20-JUN-1996
VERSION	W75342.1			
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				

```
/clone="IMAGE:390869"  
/sex="unknown"  
/tissue_type="embryo"  
/dev_stage="13.5-14.5dpc total fetus"  
/lab_host="DH10B"  
/clone_lib="Soares mouse embryo NDMR3.5 14.5"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5',  
TGTACCAATCTGAAGTGGAGCGCGGAAATTTTATTTTATTTT  
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2  
14.5dpc embryos (total RNA provided by Minoru Ko; Wayne  
State Univ., from 2 ); double-stranded cDNA was ligated to  
Eco RI adaptors (Pharmacia), digested with Not I and  
cloned into the Not I and Eco RI sites of the modified  
pT73 vector. Library went through one round of  
normalization, and was constructed by Bento Soares and  
M.Fatima Bonaldo."
```

```
BASE COUNT      101 a      141 c      123 g      97 t  
ORIGIN
```

```
Query Match      40.2%; Score 261.6; DB 14; Length 462;
```

```
Best Local Similarity 79.2%; Pred. No. 3.4e-44;  
Matches 350; Conservative 0; Mismatches 84; Indels 8; Gaps 3;
```

```
QY      1 GGAAGAGACGCGCTGGAACCATGACACGCGGTATCCCGCGAGAGACACCGGCCCC 60  
Db      5 GGAAGACACGCGCGCTGGAACCATGACATTCATATCCCGTGA-GACCCCGGCGCTCC 63  
QY      61 CAGGCCGAGACGCGCGT-----GCCACACAGCGCTCACTGGCGGCCCGCACCC 114  
Db      64 ATCATCCCGAGGCTGATGCTGCGACACAGCCCTCTCCATGGGAAACCTGGGCC 123  
QY      115 CCCGCTCGAGACCATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 174  
Db      124 TAGACACGAGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183  
QY      175 CTTGGCTTCTGCGCTGCGCTTACTCATCAGGCCGAGATCAGAAAGTGTGTGA 234  
Db      184 CTTGGATTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243  
QY      235 CTTGGAAGCGGCGCGGCTTGTGCTCAAA-GCCAGTGTACATCTGCGCGCGA 293  
Db      244 CTTGGAAGCTGCAAGGCAATGCTGCTCAAGGCCAAGTGTACATCTGCTGCA 303  
QY      294 TGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 353  
Db      304 TGTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363  
QY      354 TGGCCCGCTGCGCAAGACTCTGCGCTTCTTCAAGACCAAGTTGATGACGGGACT 413  
Db      364 TGTCCAAGTTAGCCAAAGACTCTGCGCTTCTTCAAGACCAAGTTGATGAGGAGACT 423  
QY      414 ATGACTGACAGGCTGGGTCTTG 435  
Db      424 ATTAATAAGATTCCGAGCTTG 445
```

```
Search completed: January 5, 2004, 18:04:06  
Job time : 1977 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 16:30:42 ; Search time 2818 Seconds

(without alignments)
9450.734 Million cell updates/sec

Title: US-09-914-815A-10

Perfect score: 651

Sequence: 1 ggaagagacggcgctggaac.....ccggggccctaactctgcc 651

Scoring table: IDENTITY NUC

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database :
1: gb_bai:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rdt:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_by:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	446.6	68.6	95038	9 AF015416	AF015416 Homo sapi
C 2	446.6	68.6	14383	9 AF015416	AF015416 Homo sapi
C 3	446.6	68.6	156819	2 AP006283	AP006283 Homo sapi
C 4	446.6	68.6	175416	2 AP006286	AP006286 Homo sapi
C 5	446.6	68.6	187160	9 AC138230	AC138230 Homo sapi
C 6	284.6	43.7	499	10 MMU9781	MMU9781 Mus muscu
C 7	173	26.6	238533	2 AC132968	AC132968 Rattus no
C 8	167.6	25.7	200734	2 AC109272	AC109272 Mus muscu
C 9	167.6	25.7	251206	2 AC107815	AC107815 Mus muscu
C 10	95.4	14.7	136098	2 AC006970	AC006970 Homo sapi
C 11	93.8	14.4	51891	2 AC074103	AC074103 Homo sapi
C 12	93.4	14.3	678	9 BC009696	BC009696 Homo sapi
C 13	93.4	14.3	905	6 AX337050	AX337050 Sequence
C 14	93.4	14.3	905	9 HS18D	HS18D Human 1-8D
C 15	93.2	14.3	402	9 BT006892	BT006892 Homo sapi
C 16	93.2	14.3	402	12 BT007876	BT007876 Synthetic
C 17	93.2	14.3	637	9 BC006794	BC006794 Homo sapi
C 18	93.2	14.3	642	6 AR225472	AR225472 Sequence
C 19	93.2	14.3	642	6 AX321588	AX321588 Sequence
C 20	93.2	14.3	645	9 BC022439	BC022439 Homo sapi
C 21	93.2	14.3	695	6 AX281855	AX281855 Sequence
C 22	93.2	14.3	749	9 BC008417	BC008417 Homo sapi
C 23	93.2	14.3	808	9 HS18U	HS18U Human 1-8U
C 24	91.6	14.1	37918	9 AP005232	AP005232 Homo sapi
C 25	91.6	14.1	222765	2 AC144988	AC144988 Gorilla g
C 26	91.2	14.0	378	9 BT007173	BT007173 Homo sapi
C 27	91.2	14.0	378	12 BT007946	BT007946 Synthetic
C 28	91.2	14.0	647	9 HSRNA927	HSRNA927 H. sapiens m
C 29	91.2	14.0	683	9 BC000897	BC000897 Homo sapi
C 30	91.2	14.0	842	6 AX397528	AX397528 Sequence
C 31	91.2	14.0	851	6 AX014337	AX014337 Sequence
C 32	91.2	14.0	853	6 AX302539	AX302539 Sequence
C 33	91.2	14.0	853	6 AX409461	AX409461 Sequence
C 34	91.2	14.0	853	6 HOMO27A	HOMO27A Human inter
C 35	91.2	14.0	1095	6 AX281854	AX281854 Sequence
C 36	91	14.0	245	6 AX340862	AX340862 Sequence
C 37	90.8	13.9	579	4 AF272041	AF272041 Bos tauru
C 38	90.8	13.9	107717	9 AC137055	AC137055 Homo sapi
C 39	90.8	13.9	110000	2 AC004085_1	AC004085_1 Continuation (2 of
C 40	90.8	13.9	110000	2 AC004085_2	AC004085_2 Continuation (3 of
C 41	90.2	13.9	122351	9 AC023157	AC023157 Homo sapi
C 42	89.6	13.8	905	10 BC027285	BC027285 Mus muscu
C 43	89.6	13.8	84979	9 AC116049	AC116049 Homo sapi
C 44	89.6	13.8	176104	2 AC016193	AC016193 Homo sapi
C 45	89.6	13.8	189572	2 AC084812	AC084812 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AF015416/c 95038 bp DNA linear PRI 20-AUG-1997
DEFINITION Homo sapiens chromosome 11 from 11p15.5 region, complete sequence.
ACCESSION AF015416
VERSION AF015416.1 GI:2335202
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 95038)
Evans, G.A., Bradbury, P., Brignac, S., Bumeister, R., Davie, J.,
Davies, C.J., Davis, C., English, C., Fondon, R., Franklin, F.L.,
Garner, H.R., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J.,

TITLE
JOURNAL
REFERENCE
AUTHORS

95 Kb from Four Overlapping Human Chromosome 1p15.5 Cosmids
Unpublished
(Bases 1 to 95038)
Evans, G.A., Bradbury, P., Brignac, S., Bumeister, R., Davie, J.,
Hinson, S., Megarity, C., Narayanaswamy, J., Newson, J., O'Brien, K.,
Oliver, T., Patel, P., Probst, S., Rayner, S., Scraggman, J.,
Schilling, P., Schuler, R., Syed, M., Valenzuela, D., Ward, T.,
Wilson, R. and Burbee, D.

Evans, G. A., Bredbury, P., Brignac, S., Bumeister, R., Davis, J., Davies, C. J., Davis, C., English, S., Fendton, T., Franklin, T. L., Garner, H. R., Gordon, M., Gowaty, G., Grant, O., Hahnert, L., Harris, J., Hines, S., Megarity, C., Narayanasamy, U., Newton, J., O'Brien, K., Oliver, T., Patel, P., Probst, S., Rayner, S., Schengen, J., Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T., Wilson, R. and Butree, D.

Direct Submission
Submitted (18-JUL-1997) Genome Science and Technology Center,
University of Texas Southwestern Medical Center at Dallas, 5323
Harry Hines Blvd, Dallas, TX 75355-8551, USA

FEATURES	Location/Qualifiers
source	1. .95038

repeat_region	711. .975
/rpt_family="Alu"	
repeat_region	2728. .3022
/rpt_family="Alu"	
repeat_region	3363. .3719
/rpt_family="MER1"	
repeat_region	5634. .5909
/rpt_family="Alu"	
repeat_region	complement(6720. .6959)
/rpt_family="Alu"	
repeat_region	10560. .10839
/rpt_family="Alu"	
repeat_region	complement(11832. .12455)
/rpt_family="Alu"	
repeat_region	12821. .13405
/rpt_family="Al"	
repeat_region	13718. .14314
/rpt_family="Alu"	
repeat_region	14617. .14779
/rpt_family="Alu"	
repeat_region	15396. .15695
/rpt_family="Alu"	
repeat_region	15722. .15994
/rpt_family="Alu"	
repeat_region	complement(16696. .16997)
/rpt_family="Alu"	
repeat_region	complement(18996. .19286)
/rpt_family="Alu"	
repeat_region	complement(19458. .19747)
/rpt_family="Alu"	
repeat_region	complement(19777. .20045)
/rpt_family="Alu"	
repeat_region	complement(122885. .23002)
/rpt_family="MIR"	
repeat_region	23109. .23261
/rpt_family="Alu"	
repeat_region	23392. .23481
/rpt_family="MER46"	
repeat_region	complement(123877. .22944)
/rpt_family="MIR"	
repeat_region	complement(124595. .24986)
/rpt_family="Alu"	
repeat_region	complement(125429. .26012)
/rpt_family="Alu"	
repeat_region	complement(12780. .27676)
/rpt_family="Alu"	
repeat_region	complement(129422. .29486)
/rpt_family="MER42"	

[illegible][illegible]

Query	Subject	Score	DB	Length	Pos
repeat_region	repeat_region	87125	.87256		
repeat_region	repeat_region	87350	.87893		
repeat_region	repeat_region	89103	.89401		
repeat_region	repeat_region	89726	.89837		
repeat_region	repeat_region	89888	.90390		
repeat_region	repeat_region	92187	.92187		
repeat_region	repeat_region	93951	.94017		
BASE COUNT	23552 a 24227 c 24425 g 22834 t				
ORIGIN					
Query Match	Best Local Similarity	68.6%	Score 446.6	DB 9	Length 95038
Matches 449	Conservative 0	Mismatches 4	Indels 0	Gaps 0	
199	CTCCATCAAGCCCGAGATCAAGAGGTGTGTGACTTGAAGCGCCGCGGCTTTTG3	258			
83096	CTCTCTCAAGCCCGAGATCAAGAGGTGTGTGACTTGAAGCGCCGCGGCTTTTG3	83033			
259	CTCCAAAGCCCAAGTGTCAACATCTCTGCGCCGAGTGTGAGCGCTGTCGCACTGCT	318			
83036	CTCCAAAGCCCAAGTGTCAACATCTCTGCGCCGAGTGTGAGCGCTGTCGCACTGCT	82977			
319	GCTCTTGGAGGCTGTGTGTGACTGTGTGCTGCACTGCGCCGCTGAGCGCCGCTGAG	378			
82976	GCTCTTGGAGGCTGTGTGTGACTGTGTGCTGCACTGCGCCGCTGAGCGCCGCTGAG	82917			
379	CGCCTTCTTGAAGCAACAGTTTGTATGACGCGGACTATGACTGACAGAGCTGCTGATC	438			
82916	CGCCTTCTTGAAGCAACAGTTTGTATGACGCGGACTATGACTGACAGAGCTGCTGATC	82855			
439	TGGGGCACTAGCCCGAGACATGACCCCAAGCTGTGCTGCCCTGGGGCCCAATAGTACT	498			
82856	TGGGGCACTAGCCCGAGACATGACCCCAAGCTGTGCTGCCCTGGGGCCCAATAGTACT	82797			
499	CCCCGAGGCTGTGCGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	558			
82796	CCCCGAGGCTGTGCGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	82737			
559	CCCTCCAGCCCAACATGAGGACCTTAAGGCTGAACAGTGAACCCCGGGGCTTCAACC	618			
82736	CCCTCCAGCCCAACATGAGGACCTTAAGGCTGAACAGTGAACCCCGGGGCTTCAACC	82677			
619	TAAACCGAGAGTTCCTCGGGCCCTAAGCTCTGCC	651			
82676	TAAACCGAGAGTTCCTCGGGCCCTAAGCTCTGCC	82644			
LOCUS	AC136475	143835 bp	DNA	linear	PRI 30-MAY-2003
DEFINITION	Homo sapiens chromosome 11, clone RP11-326C3, complete sequence.				
ACCESSION	AC136475				
VERSION	AC136475.6	GI:31193978			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 143835)				
AUTHORS	Bitren,B., Nussbaum,C. and Lander,E.				
TITLE	Homo sapiens chromosome 11, clone RP11-326C3				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 143835)				
AUTHORS	Bitren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Bana,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,				

Canarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gird, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, S., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (02-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 143835)

Bliren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Baetien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Canarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, S., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 143835)

Bliren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Baetien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Canarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, S., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (30-MAY-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 30, 2003 this sequence version replaced gi:2946599.

All repeats were identified using RepeatMasker:

Smith, A.P.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

DB 82847 TAACCCGAGAGTCCCGGCGCTTAAGTCTGCC 82815

RESULT 3
AP006283
LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-326C3 map 11p, WORKING DRAFT
SEQUENCE, 19 unordered pieces.
ACCESSION AP006283
VERSION AP006283.1 GI:2924343
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Homo sapiens genomic DNA of 11p
JOURNAL Published Only in Database (2003)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Shohiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
Project Information
Center project name: HumRef11
Center clone name: RP11-326C3
Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 144706 bases at least Q40
Consensus quality: 150967 bases at least Q20
Consensus quality: 153657 bases at least Q20
Insert size: 155019; sum-of-contigs
Quality coverage: 6.36x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 19 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1
21001 contig of 21001 bp in length
21102 40714 contig of 19613 bp in length
40815 58430 contig of 17616 bp in length
58531 66754 contig of 8224 bp in length
66855 81134 contig of 14280 bp in length
81235 89763 contig of 8529 bp in length
89864 101889 contig of 12026 bp in length
101990 112726 contig of 7077 bp in length
112903 125181 contig of 5279 bp in length
125282 129184 contig of 3903 bp in length
129285 137261 contig of 7976 bp in length
137361 143543 contig of 6182 bp in length
143642 146891 contig of 3249 bp in length
146992 149568 contig of 2577 bp in length
149669 152678 contig of 3010 bp in length
152779 154135 contig of 1357 bp in length
154236 155662 contig of 1427 bp in length
155763 156819 contig of 1057 bp in length

NOTE: This is a 'working draft' sequence. It currently consists of 19 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1
21001: contig of 21001 bp in length
21102: gap of 100 bp
40714: contig of 19613 bp in length
40815: gap of 100 bp
58430: contig of 17616 bp in length
58531: gap of 100 bp
66754: contig of 8224 bp in length
66855: gap of 100 bp
81134: contig of 14280 bp in length
81235: gap of 100 bp
89763: contig of 8529 bp in length
89864: gap of 100 bp
101889: contig of 12026 bp in length
101990: gap of 100 bp
112726: contig of 7077 bp in length
112903: gap of 100 bp
125181: contig of 5279 bp in length
125282: gap of 100 bp
129184: contig of 3903 bp in length
129285: gap of 100 bp
137261: contig of 7976 bp in length
137361: gap of 100 bp
143543: contig of 6182 bp in length
143642: gap of 100 bp
146891: contig of 3249 bp in length
146992: gap of 100 bp
149568: contig of 2577 bp in length
149669: gap of 100 bp
152678: contig of 3010 bp in length
152779: gap of 100 bp
154135: contig of 1357 bp in length
154236: gap of 100 bp
155662: contig of 1427 bp in length
155763: gap of 100 bp
156819: contig of 1057 bp in length.

FEATURES
Source
1..156819
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p"
/clone="RP11-326C3"
1..21001
/note="assembly_fragment"
21102..40714
/note="assembly_fragment"
40815..58430
/note="assembly_fragment"
58531..66754
/note="assembly_fragment clone_end:sp6 vector_side:left"
66855..81134
/note="assembly_fragment"
81235..89763
/note="assembly_fragment"
89864..101889
/note="assembly_fragment"
101990..112726
/note="assembly_fragment"
112726..119903
/note="assembly_fragment"
119903..125181
/note="assembly_fragment"

misc_feature
21102..40714
misc_feature
40815..58430
misc_feature
58531..66754
misc_feature
66855..81134
misc_feature
81235..89763
misc_feature
89864..101889
misc_feature
101990..112726
misc_feature
112726..119903
misc_feature
119903..125181

```

misc_feature 125282..129184
              /note="assembly_fragment_clone_end:T7_vector_side:right"
misc_feature 129285..137260
              /note="assembly_fragment"
misc_feature 137361..143542
              /note="assembly_fragment"
misc_feature 143643..146891
              /note="assembly_fragment"
misc_feature 146992..149568
              /note="assembly_fragment"
misc_feature 149669..152678
              /note="assembly_fragment"
misc_feature 152779..154135
              /note="assembly_fragment"
misc_feature 154236..155662
              /note="assembly_fragment"
misc_feature 155763..156819
              /note="assembly_fragment"
BASE COUNT 37089 a 39987 c 40186 g 37757 t 1800 others
ORIGIN

```

```

Query Match      68.6%; Score 446.6; DB 2; Length 156819;
Best Local Similarity 99.1%; Pred. No. 2.3e-78;
Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Oy 199 CTCATCAAGGCGCGAGATCAAGAGTGTGTGACCTGGAAGCGCGCGCTTTTGG 258
Db 92819 CTCCTGAGGCGCGAGATCAAGAGTGTGTGACCTGGAAGCGCGCGCTTTTGG 92878

Oy 259 CTCMAAGCAGAGTCTCAACATCTTGGCGCGAGTGTGAGCGTGTGCGCACTGCT 318
Db 92879 CTCMAAGCAGAGTCTCAACATCTTGGCGCGAGTGTGAGCGTGTGCGCACTGCT 92938

Oy 319 GCTCTGGGCGTGGTGTGATGAGTGTGACCTGACCTGGCGCGCGCAAGGACTGCG 378
Db 92939 GCTCTGGGCGTGGTGTGATGAGTGTGACCTGACCTGGCGCGCGCAAGGACTGCG 92998

Oy 379 GCGCTTCTTCAAGCAGAGTGTGATGAGCGGAGTGTGAGCGTGTGCGCTGATC 438
Db 92999 GCGCTTCTTCAAGCAGAGTGTGATGAGCGGAGTGTGAGCGTGTGCGCTGATC 93058

Oy 439 TGGGCACTAAGCCCAAGACACTGACCCAGGCTGCTGCCCTGGGCGCAATCTGACT 498
Db 93059 TGGGCACTAAGCCCAAGACACTGACCCAGGCTGCTGCCCTGGGCGCAATCTGACT 93118

Oy 499 CCGGAGAGCTGGGCGCTCTTGTGAGGCGCTCCATCCCTGCCATCTGATCTGGGG 558
Db 93119 CCGGAGAGCTGGGCGCTCTTGTGAGGCGCTCCATCCCTGCCATCTGATCTGGGG 93178

Oy 559 CCCTTCAGCCCAAGATGGGCACTTAAGCTGAACCAATCAAGACCCCGGCTCTTCAACC 618
Db 93179 CCCTTCAGCCCAAGATGGGCACTTAAGCTGAACCAATCAAGACCCCGGCTCTTCAACC 93238

Oy 619 TAACCCGAGATTCCTGGGCGCTTAATCTGCCC 651
Db 93239 TAACCCGAGATTCCTGGGCGCTTAATCTGCCC 93271

```

```

RESULT 4
AP006286/c 175416 bp DNA linear HTG 26-MAR-2003
LOCUS Homo sapiens chromosome 11 clone RP13-25N22 map 11p, WORKING DRAFT
DEFINITION
SEQUENCE, 22 unordered pieces.
ACCESSION AP006286
VERSION AP006286.1 GI:29243346
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

```

```

TITLE Homo sapiens genomic DNA of 11p
JOURNAL Published Only in Database (2003)
REFERENCE 2 (bases 1 to 175416)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
tel:81-45-503-9111, fax:81-45-503-9170)
COMMENT
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
Project Information
Center project name: Humdrat11
Center clone name: RP13-25N22

```

----- Summary Statistics

```

Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator RT-ampliphase; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 160074 bases at least Q40
Consensus quality: 167722 bases at least Q30
Consensus quality: 171662 bases at least Q20
Insert size: 173316; sum-of-contigs
Quality coverage: 7.92x in Q20 bases; sum-of-contigs

```

NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1 33131 contig of 33131 bp in length
33232 54098 contig of 20667 bp in length
54199 72037 contig of 17039 bp in length
72138 83277 contig of 11140 bp in length
83378 98599 contig of 15222 bp in length
98700 10513 contig of 6814 bp in length
105614 116229 contig of 10616 bp in length
116330 121964 contig of 5635 bp in length
12065 127627 contig of 5563 bp in length
12728 131971 contig of 4244 bp in length
132072 137506 contig of 5435 bp in length
137607 143829 contig of 6223 bp in length
143930 148539 contig of 4610 bp in length
148640 153148 contig of 4509 bp in length
153249 157509 contig of 4252 bp in length
157610 160561 contig of 2952 bp in length
160662 164778 contig of 4117 bp in length
164879 168269 contig of 3391 bp in length
168370 170689 contig of 2320 bp in length
170790 172833 contig of 2044 bp in length
172934 174155 contig of 1122 bp in length
174256 175416 contig of 1161 bp in length.
NOTE: This is a 'working draft' sequence. It currently
consists of 22 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 33131: contig of 33131 bp in length
33132 33231: gap of 100 bp
33232 54098: contig of 20667 bp in length
54099 54198: gap of 100 bp
54199 72037: contig of 17039 bp in length
72038 72137: gap of 100 bp
72138 83277: contig of 11140 bp in length

```

```

*      83378      83377: gap of 100 bp
*      83378      98599: contig of 15222 bp in length
*      98600      98699: gap of 100 bp
*      98700      10513: contig of 6814 bp in length
*      105614     105613: gap of 100 bp
*      105614     116229: contig of 10616 bp in length
*      116330     116329: gap of 100 bp
*      116330     121964: contig of 5635 bp in length
*      121965     122064: gap of 100 bp
*      122065     127627: contig of 5563 bp in length
*      127628     131971: contig of 4244 bp in length
*      131972     132071: gap of 100 bp
*      132072     137506: contig of 5435 bp in length
*      137507     137606: gap of 100 bp
*      137607     143829: contig of 6223 bp in length
*      143830     143929: gap of 100 bp
*      143930     148539: contig of 4610 bp in length
*      148540     148639: gap of 100 bp
*      148640     153148: contig of 4509 bp in length
*      153149     153248: gap of 100 bp
*      153249     157509: contig of 4261 bp in length
*      157510     157609: gap of 100 bp
*      157610     160561: contig of 2952 bp in length
*      160562     160661: gap of 100 bp
*      160662     164778: contig of 4117 bp in length
*      164779     164878: gap of 100 bp
*      164879     168269: contig of 3391 bp in length
*      168270     168369: gap of 100 bp
*      168370     170689: contig of 2320 bp in length
*      170690     170789: gap of 100 bp
*      170790     172833: contig of 2044 bp in length
*      172834     172933: gap of 100 bp
*      172934     174155: contig of 1222 bp in length
*      174156     174255: gap of 100 bp
*      174256     175416: contig of 1161 bp in length.

```

FEATURES

source

```

1. 175416
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="11"
   /map="11p"
   /clone="RP13-25N22"
   1. 33131
   /note="assembly_fragment clone_end:SP6 vector_side:left"
   misc_feature
   33332..54098
   /note="assembly_fragment"
   misc_feature
   54199..72037
   /note="assembly_fragment"
   misc_feature
   72138..83277
   /note="assembly_fragment"
   misc_feature
   83378..98599
   /note="assembly_fragment"
   misc_feature
   98700..105513
   /note="assembly_fragment clone_end:T7 vector_side:left"
   misc_feature
   105614..116229
   /note="assembly_fragment"
   misc_feature
   116330..121964
   /note="assembly_fragment"
   misc_feature
   122065..127627
   /note="assembly_fragment"
   misc_feature
   127728..131971
   /note="assembly_fragment"
   misc_feature
   132072..137506
   /note="assembly_fragment"
   misc_feature
   137607..143829
   /note="assembly_fragment"
   misc_feature
   143930..148539
   /note="assembly_fragment"
   misc_feature
   148640..153148
   /note="assembly_fragment"
   misc_feature
   153249..157509

```

```

/note="assembly_fragment"
157610..160561
/note="assembly_fragment"
160662..164778
/note="assembly_fragment"
164879..168269
/note="assembly_fragment"
168370..170689
/note="assembly_fragment"
170790..172833
/note="assembly_fragment"
172934..174155
/note="assembly_fragment"
174256..175416
/note="assembly_fragment"
BASE COUNT   37167 a 49392 c 49417 g 37339 t 2101 others
ORIGIN
Query Match      68.6%; Score 446.6; DB 2; Length 175416;
Best Local Similarity 99.1%; Pred. No. 2,2e-78;
Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      199      CTCATCAAGGCGCCGAGATCAGAGGTGCTGTAAGCGGCGCGGCTTTGG 258
DB      42209      CTCCTGCAAGCCCGAGATCAGAGGTGCTGTAAGCGGCGCGGCTTTGG 42150
QY      259      CTCGAAGCGCAATGCTACACATCTCTGGCCGAGATGTGACGCTGTCGCACTGCT 318
DB      42149      CTCGAAGCGCAATGCTACACATCTCTGGCCGAGATGTGACGCTGTCGCACTGCT 42090
QY      319      GCTCTGGGGCTGTGTGTGATCTGTGCTGACCTGACCTGGCCCGGCTGGCCAGGACTCTGC 378
DB      42089      GCTCTGGGGCTGTGTGTGATCTGTGCTGACCTGACCTGGCCCGGCTGGCCAGGACTCTGC 42030
QY      379      CGCCTCTTCAGACCAAGTTGATATACGCGGACTATGACTGACAGGCTGGCTTCGATC 438
DB      42029      CGCCTCTTCAGACCAAGTTGATATACGCGGACTATGACTGACAGGCTGGCTTCGATC 41970
QY      439      TGGGGCACTAGCCCGCAGACACTGACCCGAGCTGCTGCCCTGGGGCCCAATGACT 498
DB      41969      TGGGGCACTAGCCCGCAGACACTGACCCGAGCTGCTGCCCTGGGGCCCAATGACT 41910
QY      499      CCCCGAGCTGGGCTCTCTCTGTGTGGGCTTCATCTCCGCTGCTGATCTGGGG 558
DB      41909      CCCCGAGCTGGGCTCTCTCTGTGTGGGCTTCATCTCCGCTGCTGATCTGGGG 41850
QY      559      CCTTCAGGCCCAACAATGGGCACTTAAGCTGAACCAAGCAAGCCCGGGGCTTTACCC 618
DB      41849      CCTTCAGGCCCAACAATGGGCACTTAAGCTGAACCAAGCAAGCCCGGGGCTTTACCC 41790
QY      619      TAACCCGAGAGTCCCGGCGCTTAAGCTGACC 651
DB      41789      TAACCCGAGAGTCCCGGCGCTTAAGCTGACC 41757

```

```

RESULT 5
AC138230/c      187160 bp      DNA      linear      PRI 22-MAY-2003
LOCUS      Homo sapiens chromosome 11, clone RP13-317D12, complete sequence.
DEFINITION      AC138230
ACCESSION      AC138230.5 GI:30984765
VERSION      HTG.
KEYWORDS      Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 187160)
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      Birren,B., Nussbaum,C. and Lander,B.
JOURNAL      Homo sapiens chromosome 11, clone RP13-317D12
REFERENCE      Unpublished
2 (bases 1 to 187160)
AUTHORS      Birren,B., Nussbaum,C., Lander,B., Ali,A., Allen,N., Anderson,S.,
Batra,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,

```

TITLE
JOURNAL
REFERENCE
AUTHORS
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Haez, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,
Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T.,
Menga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N.,
Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K.,
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS
Submitted (20-DEC-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 187160)
Birren, B., Nussbaum, C., Lander, E., Abouelella, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cooke, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Haez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Meneus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Strange-Thomann, N., Stojanovic, N., Strubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 187160)
Birren, B., Nussbaum, C., Lander, E., Abouelella, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cooke, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Haez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Meneus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Strange-Thomann, N., Stojanovic, N., Strubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
Web site: <http://www-seg.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: l29013
Center clone name: 317_D12

FEATURES
source
Location/Qualifiers
1.187160
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP13-317D12"
/clone_1b="RP13-317D12 Human Female BAC"
135..200
repeat_region
/rpt_family="L2"
1134..1310
repeat_region
/rpt_family="WIR"
complement(1311..1614)
/rpt_family="AluY"
1615..1638
repeat_region
/rpt_family="WIR"
1849..1938
repeat_region
/rpt_family="L1ME"
complement(14297..4385)
/rpt_family="L2"
4463..4561
repeat_region
/rpt_family="FRAM"
complement(4573..4613)
/rpt_family="WAD1"
4733..5040
repeat_region
/rpt_family="AluY"
complement(5094..5383)
/rpt_family="L2"
complement(5379..5505)
/rpt_family="L2"
complement(5501..5627)
/rpt_family="L2"
complement(5667..6096)
/rpt_family="L2"
6542..6666
repeat_region
/rpt_family="MER41E"
6699..6719
repeat_region
/rpt_family="GC_rich"
6917..6996
repeat_region
/rpt_family="C-rich"
7136..7329
repeat_region
/rpt_family="C-rich"
15461..15483
repeat_region
/rpt_family="TG)n"
18178..18413
repeat_region
/rpt_family="L2"
complement(18543..18907)
/rpt_family="MER4A"
complement(18913..19138)
/rpt_family="AluSg/x"
complement(19139..19318)
/rpt_family="MER4D"
complement(19319..19688)
/rpt_family="MER4B1"
19689..19992
repeat_region
/rpt_family="AluSp"
19994..20308
repeat_region
/rpt_family="AluSg"
complement(20309..20630)
/rpt_family="MER4E1"
20632..20763
repeat_region
/rpt_family="AluSg/x"
complement(20785..20882)
/rpt_family="MER4E"
complement(20886..21091)

```

repeat_region      /rpt_family="MER41E"
21092..21402
/rpt_family="AlusX"
repeat_region      complement(21403..21580)
/rpt_family="MER41E"
repeat_region      complement(22261..22538)
/rpt_family="MER41E"
repeat_region      complement(22543..22594)
/rpt_family="MER41E"
repeat_region      complement(22595..22902)
/rpt_family="AluY"
repeat_region      complement(22903..23195)
/rpt_family="MER4D"
repeat_region      complement(23197..23314)
/rpt_family="FLAM_C"
repeat_region      complement(23318..23368)
/rpt_family="MER4D"
repeat_region      complement(23369..23890)
/rpt_family="MER4A"
repeat_region      complement(23891..24166)
/rpt_family="MER4D"
repeat_region      complement(24381..24698)
/rpt_family="MER63"
repeat_region      complement(25372..25668)
/rpt_family="AluY"
repeat_region      complement(26402..26493)
/rpt_family="(TG)n"
repeat_region      complement(27325..28044)
/rpt_family="LTR30"
unsure             /note="<30 qual SNGL region"
repeat_region      29024..29047
/rpt_family="GC_rich"
repeat_region      complement(29397..30103)
/rpt_family="LTR8"
repeat_region      30537..30656
/rpt_family="LTR2C"
repeat_region      complement(30657..31083)

Query Match      68.6%; Score 446.6; DB 9; Length 187160;
Best Local Similarity 99.1%; Pred. No. 2.2e-78;
Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 199 CTCCTCAACAGGCCCGGAGATCAAGAGTGTGTGTGACCTGGAACCGCGCGCTTTTGG 258
DB 16606 CTCCTCAACAGGCCCGGAGATCAAGAGTGTGTGTGACCTGGAACCGCGCGCTTTTGG 16547
QY 259 CTCCTCAACAGGCCCGGAGATCAAGAGTGTGTGTGACCTGGAACCGCGCGCTTTTGG 318
DB 16546 CTCCTCAACAGGCCCGGAGATCAAGAGTGTGTGTGACCTGGAACCGCGCGCTTTTGG 16487
QY 319 GCTCTGCGGGCTGTGTGTGACCTGTGACCTGGAACCGCGCGCTTTTGG 378
DB 16486 GCTCTGCGGGCTGTGTGTGACCTGTGACCTGGAACCGCGCGCTTTTGG 16427
QY 379 CGCTCTTTCAGACCAAGTTTATGACCGGACCTTATGACCTGGAACCGCGCGCTTTTGG 438
DB 16426 CGCTCTTTCAGACCAAGTTTATGACCGGACCTTATGACCTGGAACCGCGCGCTTTTGG 16367
QY 439 TGGGGGACTAGCCCGGAGACCTGACCCGAGCTGTGCGCCCGGAGCCCAATACGACT 498
DB 16366 TGGGGGACTAGCCCGGAGACCTGACCCGAGCTGTGCGCCCGGAGCCCAATACGACT 16307
QY 499 CCGCGAGGCTGAGCCCTCTTCTGTGTGGGCTTCATCCCTGCGCCCATCTGATCTGGGG 558
DB 16306 CCGCGAGGCTGAGCCCTCTTCTGTGTGGGCTTCATCCCTGCGCCCATCTGATCTGGGG 16247
QY 559 CCTTCAAGCCCAACAGTGGGACCTTAAAGCTGAACCAAGTCAAGCCCGGAGCTTTTCA 618
DB 16246 CCTTCAAGCCCAACAGTGGGACCTTAAAGCTGAACCAAGTCAAGCCCGGAGCTTTTCA 16187
QY 619 TAACCCGAGAGTTCGGGGCCCTTACTGTGCC 651

```

```

DB 16186 TAACCCGAGAGTTCGGGGCCCTTACTGTGCC 16154

RESULT 6
MMU9781      499 bp  mRNA  linear  ROD 22-MAR-2001
LOCUS      Mus musculus mRNA for putative haemopoietic membrane protein.
DEFINITION
VERSION      AJ009781.1 GI:3355654
KEYWORDS      haemopoietic membrane protein; putative.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS      Baird,J.W., Ryan,K.M., Hayes,I., Hampson,L., Heyworth,C.M.,
Clark,A., Mooton,M., Ansell,J.D., Menzel,U., Hole,N. and
Graham,G.J.
TITLE      Differentiating embryonal stem cells are a rich source of
JOURNAL      haemopoietic gene products and suggest erythroid preconditioning of
MEDLINE      primitive haemopoietic stem cells
PUBMED      J. Biol. Chem. 276 (12), 9189-9198 (2001)
21153738
PUBMED      11106657
2 (bases 1 to 499)
REFERENCE
AUTHORS      Graham,G.J.
TITLE      Direct Submision
JOURNAL      Submitted (24-JUL-1998) Graham G.J., CRC Beatson Laboratories,
Beatson Institute for Cancer Research, Gartnavel Estate, Switchback
Road, Bearsden, Glasgow G61 1BD, SCOTLAND, UK

FEATURES
source
1..499
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cclone="BRC1"
/cell_line="ES"
/dev_stage="d5 of in vitro development"
31..436
/note="ORF1"
/codon_start=1
/product="putative haemopoietic membrane protein"
/protein_id="CA08829.1"
/db_xref="GI:3355655"
/transcript="MDTSPREDPAPSRKADAATLMSGTPTPTDMLMSVF
STMVNLCCIGFLAVHSVKARDQDGLAARQYGSAAKCYKNTILAAWMTLVLPILL
LDIVVTGALHLKSKAKDAAPSTPTPDEBDYN"

BASE COUNT      107 a 155 c 137 g 100 t

Query Match      43.7%; Score 284.6; DB 10; Length 499;
Best Local Similarity 79.6%; Pred. No. 4.8e-46;
Matches 351; Conservative 0; Mismatches 84; Indels 6; Gaps 1;

QY 1 GGAAGAGAGCGGCTGGAACCCATGACACGCGATATCCCGGAGACACCGGAGCC 60
DB 9 GGAAGAGAGCGGCTGGAACCCATGACACCTTATATCCCGGAGACACCGGAGCC 68
QY 61 CACGCCGACGAGGCCGCTG-----GCCACACACCCCTCAACATGCGGCGCCGAC 114
DB 69 ATATATCCGCAAGGCTGATGCTGACGACCAACACCCCTCAATGAGGAACCTGAG 128
QY 115 CCGCGCTGAGACCACTTATGCTGCTGTGTGACACCTCTAATCTGATGTTG 174
DB 129 TACACCAAGAGTCAATGCTGCTGTGTGTGACACCACTGATGATGCTG 188
QY 175 CCTGCGCTTCTGCGGCTGCTGCTCAATCAAGCCCGGAGATCAAGAGTGTGTGA 234
DB 189 CTTGATTCCTGCGGCTGCTGCTCACTGTGTCAAGCCCGGAGACCAAGATGCTGGA 248
QY 225 CTGGAAGAGCGCGGCTTGTGCTTCAAGCCCAAGGCTCAATCTGCGCGCAT 294

```

[illegible]

```

TITLE
JOURNAL
REFERENCE
AUTHORS
YU,F., ZHANG,J., ZHOU,J., ZHOU,S., ZHAO,S., DUAN,D., VON
NEIDENHAEUSEN,A., WEISS,R., SMITH,D.R., HOLT,R.A., SMITH,H.O.,
WEINSTECK,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 238533)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (05-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 238533)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:22264838.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly ('a','contigs-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: KBMS
Center clone name: CH230-172C3
----- Summary Statistics
Assembly program: Phrap; Version 0.990329
Consensus quality: 223317 bases at least Q40
Consensus quality: 225463 bases at least Q30
Consensus quality: 226655 bases at least Q20
Estimated insert size: 230654; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* [see http://www.hgsc.bcm.tmc.edu/docs/Gendraft\_data.html] .
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 222461: contig of 222461 bp in length
* 222462 222561: gap of unknown length
* 222562 224949: contig of 2388 bp in length
* 224950 225049: gap of unknown length
* 225050 228242: contig of 3193 bp in length
* 228243 228342: gap of unknown length
* 228343 238533: contig of 10191 bp in length.
Location/Qualifiers
1..238533
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-172C3"
1..2349
/note="wgs_contigs"
misc_feature
source
FEATURES
COMMENT
REFERENCE
AUTHORS
JOURNAL
TITLE

```

```

misc_feature      124155..209379
                  /note="clone_boundary
                  site="end:T7
                  end_sequence:BH290527"
BASE COUNT      61451 a 53521 c 52085 g 60005 t 11471 others
ORIGIN
Query Match      26.6%; Score 173; DB 2; Length 238533;
Best Local Similarity 81.6%; Pred. No. 1,4e-24;
Matches 200; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 201 CCATCAAGGCCGAGATCAGAAAGGTTGTGACCTTGGAAGCGGCCGGCTTTGGCT 260
DB 213892 CCTCGAGGCCCGAGACAGAGAGTGAAGCCGGGAGCTTGAGAGCTGCAAGAGATGAGCT 213833
QY 261 CCAAGCCAAAGTGTATCAACATCTCTGCGCGGATGTGACGCTGTGCGCCACTGTGTC 320
DB 213832 CCAAGCCAAAGTGTATCAACATCTCTGCGCGGATGTGACGCTGTGCGCCACTGTGTC 213773
QY 321 TCCCTGGGCTGTGTGTGATCTGTGCTCTGCACTTGCGCCGCTGCGCCAGAGACTGTGCG 380
DB 213772 TCCCTGGGCTGTGTGTGATCTGTGCTCTGCACTTGCGCCGCTGCGCCAGAGACTGTGCG 213713
QY 381 CCTCTTCAGCAGCCAAAGTTGATGACGCGGACTATGACTGACAGGCTGGGTCTGATCTG 440
DB 213712 CTTCTTCAGCAGCCAAAGTTGATGAGAGACTATGACTGAGAGTTCTGACCTGTGCTG 213653
QY 441 GAGCA 445
DB 213652 AACCA 213648

RESULT 8
ACIO9272/c      200734 bp DNA linear HTG 23-APR-2003
LOCUS           Mus musculus clone RP23-354L18, WORKING DRAFT SEQUENCE, 17 ordered
DEFINITION      pieces.
ACCESSION      ACIO9272
VERSION        ACIO9272.3 GI:3018119
KEYWORDS       HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 200734)
AUTHORS        Birren,B., Nusbaum,C. and Lander,E.
JOURNAL        Mus musculus, clone RP23-354L18
TITLES         Unpublished
                2 (bases 1 to 200734)
AUTHORS        Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
                Anderson,S., Barna,N., Baetien,V., Boguslavsky,I., Boukhalter,B.,
                Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,J.,
                Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
                Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
                Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
                Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
                Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
                Kamat,A., Karatas,A., Kells,C., Larocque,K., Lamazares,R.,
                Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
                McDonald,P., Major,J., Margulis,N., Matthews,C., McCarthy,M.,
                McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
                Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
                Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
                Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
                Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
                Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S.,
                Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
                Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
                Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
                Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
                Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
                Direct Submission
TITLES         Direct Submission

```

```

JOURNAL
REFERENCE
AUTHORS
Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 200734)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Baetien,V., Bloom,T.,
Boguslavsky,I., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Grand-Pierre,N., Hagos,B., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Liu,A., Mabbitt,R., Maclean,C.,
McDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 17, 2003 this sequence version replaced gh1:20330961.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RN/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L20944
Center clone name: 354_L_18

NOTE: This is a 'working draft' sequence. It currently
consists of 17 contigs. Gaps between the contigs
are represented as runs of 'N'. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that have
been provided by the submitter.
This sequence will be replaced
by the finished sequence as soon as it is available and
the accession number will be preserved.
1 5430: contig of 5430 bp in length
5431 5530: gap of 100 bp
5531 6612: contig of 1082 bp in length
6613 6712: gap of 100 bp
6713 8249: contig of 1537 bp in length
8250 8349: gap of 100 bp
8350 10942: contig of 2593 bp in length
10943 11042: gap of 100 bp
11042 14087: contig of 3045 bp in length
14088 14187: gap of 100 bp
14188 20506: contig of 6319 bp in length
20507 20606: gap of 100 bp
20607 26335: contig of 5729 bp in length
26336 26435: gap of 100 bp
26436 31423: contig of 4988 bp in length
31424 31523: gap of 100 bp
31524 39064: contig of 7541 bp in length
39065 39164: gap of 100 bp
39165 47880: contig of 8766 bp in length
47891 47990: gap of 100 bp
47991 56258: contig of 8268 bp in length
56259 56358: gap of 100 bp
66903: contig of 10545 bp in length
67003: gap of 100 bp
109182: contig of 42179 bp in length

```


TITLE
JOURNAL
COMMENT

Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Meneses, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Nordu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkharn, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Teffaye, S., Theodore, U., Topnam, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wymann, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Submitted (06-JUN-2003). Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 6, 2003 this sequence version replaced g1:20163115.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: 120545

Center clone name: 114_A6

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 250046 bases at least Q40

Consensus quality: 250590 bases at least Q20

Consensus quality: 250806 bases at least Q20

Insert size: 230000; agarose-1p

Insert size: 250906; sum-of-coverage

Quality coverage: 12.6 in Q20 bases; agarose-1p

Quality coverage: 11.6 in Q20 bases; sum-of-coverage

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 29459: contig of 29459 bp in length
* 29460 29559: gap of 100 bp
* 29560 30528: contig of 969 bp in length
* 30529 30628: gap of 100 bp
* 30629 142773: contig of 112145 bp in length
* 142774 142873: gap of 100 bp
* 142874 251206: contig of 108333 bp in length.
* Location/Qualifiers
1. 251206

FEATURES

source

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone_id="RP23-114A6"

/clone_11b="RP23-114A6"

1. 29459

/note="assembly_fragment"

clone_end:SP6

vector_side:left"

29560..30528

/note="assembly_fragment"

30629..142773

/note="assembly_fragment"

142874..251206

/note="assembly_fragment"

clone_end:T7

vector_side:right"

BASE COUNT 67395 a 56774 c 57538 g 69199 t 300 others

ORIGIN

Query Match 25.7% Score 167.6; DB 2; Length 251206;
Best Local Similarity 63.8%; Pred. No. 1,66-23;
Matches 277; Conservative 0; Mismatches 144; Indels 13; Gaps 1;

201 CCATCAAGGCGCCGAGATGCAAGAGGTGTGTGACCTTGAAGCGGCCGCTTTTGCT 260
131688 CCTTCAGAGCCCGAGACAGAGAGATGCTGGGAACCTTGAAGGCTCAAGGAGTATGGCT 131629
261 CCAAGCCCAAGTGTCTCAACAACATCTCGGCCGAGATGTGACGCTGTGCCCACTGTCC 320
131628 CCAAGCCCAAGTGTCTCAACAACATCTCGGCCGAGATGTGACGCTGTGCCCACTGTCC 131569
321 TCTTGAGGCTGTGTGTGTGACGCTGTGCCGCTGTGACCTGTGACGCTGTGACGCTGTG 380
131568 TCTTGAGGCTGTGTGTGTGACGCTGTGCCGCTGTGACCTGTGACGCTGTGACGCTGTG 131509
381 CTTTCTTCAAGCAACAGTTTATGACGCGGACCTATGACGACGCTGTGATCTGTG 440
131508 CTTTCTTCAAGCAACAGTTTATGACGCGGACCTATGACGACGCTGTGATCTGTG 131454
441 GGGCAGTACGCGCCAGAGACACGACCCGAGGCTGTGCCCTGTGCCCAATATGACTGCC 500
131453 -----TCCCTGAACCGAGGACACACCATGATGAGTGTGCTGTGCCCAACAGCTGC 131402
501 CCGAGCCTGAGCCTCTCTTGTGTGTGAGGCTGTGACCTGTGACCTGTGATCTGTGAGG 560
131401 TGGAGGTATGACGCTATGATGACGCTGTGATCTGTGACCTGTGACCTGTGAGGCGG 131342
561 CTCAGGCCCCCAATGGGACCTTAAAGGCTGAACGATGACGACCCGGGGGTCTTCAACCTTA 620
131341 CCTCTTGACAGATGTATTTATGATGTGCTGCCACATGATGATGATGATGATGATGAT 131282
621 ACCGAGAGTGTCC 634
131281 AACCTAAGGAGTTC 131268

RESULT 10

AC006970

LOCUS

DEFINITION

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970

AC006970 136098 bp DNA linear PRI 27-SEP-2000
Homo sapiens clone RP4-725G10, complete sequence.
AC006970.6 GI:10312289
HTG.
SOURCE
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 136098)
Waterson, R.H.
The sequence of Homo sapiens clone
unpublished
2 (bases 1 to 136098)
Waterson, R.H.
Direct Submission
Submitted (05-MAR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
3 (bases 1 to 136098)
Waterson, R.H.
Direct Submission
Submitted (27-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Sep 27, 2000 this sequence version replaced g1:9838025.
Center project name: H.DU0725G10.
Location/Qualifiers
1. 136098
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: <http://image.jnl.gov>
Series: IRAC Plate: 14 Row: a Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 10835237.

```

source
location/Qualifiers
1..678
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocustID:10581"
/db_xref="taxon:9606"
/clone="MGC:9196 IMAGE:3876542"
/tissue_type="lung, large cell carcinoma"
/clone_fib="NIH MGC_68"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
31..429

```

BASE COUNT	146 a	231 c	148 g	153 t
ORIGIN				

	Query Match	Similarity	62.8%	Score 93.4	DB 9	Length 678
	Best Local	Similarity	62.8%	Pred. No. 1.8e-08		
	Matches	145	Conservative	0	Mismatches 86	Indels 0
					Gaps	0
QY	125	GACCACTTGAATCTGATGCTGCTGTTTCAGCAACCTCTACCTGAATCTGTGTGCTCGCTGCCTTC	184			
Db	193	GACCATATGCTGCTCTGTGCTGCTGCTGTTCAACACCTCTTCATGAAACAACCTGCTGCTGGGCTTC	252			
QY	185	CTGGCGCTGCGCTTCTCCATCATGAAGCGCCGAGATCAGAGAGTGTGTGTGACTTGGAAACG	244			
Db	253	ATAGCATTCGCTTACTCTCCGTAAAGTCTAAGGACAGAAAGATGTTGGAGAGTACCTGGG	312			
QY	245	GCCCCGGGCTTTTGGCTCCAAAGCCAGTGTACACATCTCTGCGCCGATGTGACGCTG	304			
Db	313	GCCCCAGGCGCTATGCTCTCCACCGCCAGATGCTGAAACATCTGGGCGCTATTTTGGGCAATC	372			
QY	305	GTCGCCGACATCTGCTCTCTCTGGGGCTGATGTGTGACTGTGTGCTGACACTG	355			
Db	373	TTCAATGACATTTCTGCTGTATCATCTCCAGATGTTGTGTCTGCACAGGCCAG	423			

RESULT	13
AX337050	
LOCUS	AX337050
DEFINITION	Sequence from Patent WO0194629.
ACCESSION	AX337050
VERSION	AX337050.1 GI:18127769
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1
AUTHORS	Young,P.B., Augustus,M., Carter,K.C., Ebner,R., Endress,G., Horligan,S., Soppet,D.R. and Weaver,Z. Cancer gene determination and therapeutic screening using signature gene sets
TITLE	Patent: WO 0194629-A 7559 13-DEC-2001;
JOURNAL	Avalon Pharmaceuticals (US) Location/Qualifiers
FEATURES	

source	1. .905			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
BASE COUNT	189 a	304 c	206 g	206 t
ORIGIN				

Query Match	14.3%;	Score 93.4;	DB 6;	length 905;
Best Local Similarity	62.8%;	Pred. No. 1.7e-08;		
Matches 145; Conservative	0;	Mismatches 86;	Indels 0;	Gaps 0

OY	125	GACCACTGATCTGTGTGGTGTAGACACCTCTAACCTGAATCTGTGTGCTCGGCTTC	184
Db	442	GACATGTGCTGTGTGCTCTTTGAACACCTCTTACGAAACACTGTGCTCGGCTTC	501
OY	185	CTGGCGCTGCGCTTACTCCATCAGAGCCCGAGATAGAAAGTGTGTGTGACTGTGAAACG	244
Db	502	ATACCATTTGCGCTTACTCCGTAACTTAGAGGACAGAAAGTGTGTGCGAGTGAACGG	561
OY	245	GCCCGGGGTTTGTGCTCCAAAGCCAAATGTCTACAAACATCTCTGGCCGCGATGTGACGCTG	304
Db	562	GCCCAAGGCTATGTGCTCCACGCGCCAAATGTGCTGAACATCTGGGGCCCTGATTTTGGGCATC	621
OY	305	GTGCGGCACTGTGCTCTCTGGGGCTGATGTGACTGTGCTGCGTCAACCTG	355
Db	622	TTTATGACCAATTCGTCTGTGCATATATCCAGATGTTGTCTGTCCAGAGCCACG	672

RESULT 14			
LOCUS	HS18D	905 bp	linear
DEFINITION	Human 1-8D gene from interferon-inducible gene family.		
ACCESSION	X57351		
VERSION	X57351.1	GI:311373	
KEYWORDS	1-8 gene family; 1-8D gene; interferon inducible gene.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		

FEATURES	source	location/Qualifiers
1..905		
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/tissue_type="lymphoid"	
	/clone_lib="loxistb"	
gene	1..905	
	/gene="1-8D"	
	189..202	
	/gene="1-8D"	
	/note="ISIR (interferon-stimulable response element)"	
	/evidence=experimental	
	280..678	
CDS	/gene="1-8D"	
	/codon_start=1	
	/protein_id="CAA40625.1"	
	/db_xref="GI:23396"	
	/db_xref="SWISS-PROT:Q01629"	

```

exon
/translaction="NMHIVQTFSPVNSGPPNEMLEKEBEVAVLGGPHNPAPPTSTV
IHIRETSVDPHVWVSLFNTLFPNMCIGFIAPFAYSVKSRDKWGDVTAQAVASTA
KCLNIWALIGIFMTLIVLIVIPVLVQAYG"
<280>.626
/gene="1-8D"
/number=1
/evidence=experimental
exon
627..2678
/gene="1-8D"
/number=2
/evidence=experimental
repeat_unit
698..765
/gene="1-8D"
/evidence=experimental
repeat_unit
766..833
/rpt_type=TANDEM
/evidence=experimental
repeat_unit
886..891
/rpt_type=TANDEM
/evidence=experimental
polya_signal
886..891
/gene="1-8D"
BASE COUNT 189 a 304 c 206 g 206 t
ORIGIN
Query Match 14.3%; Score 93.4; DB 9; Length 905;
Best Local Similarity 62.8%; Pred. No. 1.7e-08;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 125 GACCACTTGATCTGGTGGTGTGACGACCCCTACCTGATCTGTGTGCTGGCTTC 184
DB 442 GACCATGTCGTGTGTCCTGCTGTTCAACACCTCTTCATGAACACCTGCTGCTGCTTC 501
QY 185 CTGGCGCTGGCTTACCTCATCAAGCCCGAGATCAAGAGGTGTGTGACCTGGAACG 244
DB 502 ATAGCATTCGCTACCTCGTGAAGTCAAGGAGAGATGTGGGACGTGACCGGG 561
QY 245 GCCCGCGCTTTGGCTCCAAAGCCAGTGTACACATCTGGCGCGAGTGTGACGCTG 304
DB 562 GCCCAGGCTATGCTCCACCGCCAGAGTGTGGAACATCTGGGCTGATTTTGGGATC 621
QY 305 GTGGCGGCACTGCTCTCTGGGCTGTGTGATCTGTGCTGCTGACCTG 355
DB 622 TTCTATGACCATTTCTGCTGTCATCATCCAGTGTGTGCTGCTGACGCTG 672
RESULT 15
BT006892 402 bp mRNA linear PRI 13-MAY-2003
LOCUS Homo sapiens interferon induced transmembrane protein 3 (1-8U)
DEFINITION mRNA, complete cds.
ACCESSION BT006892
VERSION BT006892.1 GI:30582622
KEYWORDS FLI CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 402)
AUTHORS Kalline,N., Chen,X., Rolfe,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
Pheasant,M., and Farmer,A.
TITLE Cloning of human full-length cDNAs in BD Creator (TM) System Donor
vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 402)
AUTHORS Kalline,N., Chen,X., Rolfe,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
Pheasant,M., and Farmer,A.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
Circle, Palo Alto, CA 94303, USA
COMMENT This cDNA clone is a part of a collection of human full length
expression clones generated by BD Biosciences Clontech and the

```

```

FEATURES
source
1..402
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GH00117X1.0"
/collection="BD Creator (TM) CDS library derived from MGC
collection"
/lab_host="DH5alpha T1 resistant"
/notes="Vector: pDNR-Dual"
1..402
/codon_start=1
/product="interferon induced transmembrane protein 3
(1-8U)"
/protein_id="AAP35538.1"
/db_xref="GI:30582623"
/translaction="NMHIVQTFSPVNSGPPNEMLEKEBEVAVLGGPHNPAPPTST
VHIRETSVDPHVWVSLFNTLFPNMCIGFIAPFAYSVKSRDKWGDVTAQAVAST
AKCLNIWALIGIFMTLIVLIVIPVLVQAYG"
BASE COUNT 77 a 140 c 98 g 87 t
ORIGIN
Query Match 14.3%; Score 93.2; DB 9; Length 402;
Best Local Similarity 66.3%; Pred. No. 2.2e-08;
Matches 134; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 125 GACCACTTGATCTGGTGGTGTGACGACCCCTACCTGATCTGTGTGCTGGCTTC 184
DB 166 GACCATGTCGTGTGTCCTGCTGTTCAACACCTCTTCATGAACACCTGCTGCTGCTTC 225
QY 185 CTGGCGCTGGCTTACCTCATCAAGCCCGAGATCAAGAGGTGTGTGACCTGGAACG 244
DB 226 ATAGCATTCGCTACCTCGTGAAGTCAAGGAGAGATGTGGGACGTGACCGGG 285
QY 245 GCCCGCGCTTTGGCTCCAAAGCCAGTGTACACATCTGGCGCGAGTGTGACGCTG 304
DB 286 GCCCAGGCTATGCTCCACCGCCAGAGTGTGGAACATCTGGGCTGATTTTGGGATC 345
QY 305 GTGGCGGCACTGCTCTCTGGGCTGTGTGATCTGTGCTGCTGACCTG 326
DB 346 CTCATGACCATTTCTGCTCATCG 367
Search completed: January 5, 2004, 17:31:03
Job time : 2823 secs

```



```

PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 17479
LENGTH: 326
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006970.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 16
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 48
OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 88
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 53
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1e+02
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 19
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 88
OTHER INFORMATION: EST HUMAN HIT: AA527771.1, EVALUATE 0.00e+00
OTHER INFORMATION: NT HIT: g111995467, EVALUATE 1.00e-129
OTHER INFORMATION: SWISSPROT HIT: Q01629, EVALUATE 3.00e-35
US-09-864-761-17479

Query Match 14.7%; Score 95.4; DB 9; Length 326;
Best Local Similarity 67.2%; Pred. No. 1.5e-15;
Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 126 ACCACTGATCTGTCGCTGCTTCCAGACCCCTTACCTGTAATCTGTGCTGGCTTCC 185
DB 209 ACCATGTTGCTGTGCTGCTTCCCTTCAACACCTCTTCAAGACCCCTGCTGCTGCTTCA 150
QY 186 TGGCGCTGCGCTTACTCATCAAGCCCGAATCAGAAAGTGTGTGATGACCTGGAACCGG 245
DB 149 TAGGTTTACCTTACTCATGAAGCTTAGGAGACAGAAAGATGTTGGCGACCTGACCGGG 90
QY 246 CCGCGCTTGTGCTCCAAAGCGAGTGTACAACTCCGCGCGATGTGACGCTGG 305
DB 89 CCGAGGCTTAGCTTCCACCGCCAAAGTGTGAACATCTGGGCTTGTGAGCATCC 30
QY 306 TGGCGCCACTGCTGCTCTGG 326
DB 29 TCATGACCATTTCTGCTCATCG 9

RESULT 2
US-09-864-761-692/c
Sequence 692, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Jemica-X-1
CURRENT FILING DATE: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456

```

```

PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263.6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 692
LENGTH: 459
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006970.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 16
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 48
OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 88
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 53
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1e+02
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 19
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 88
US-09-864-761-692

Query Match 14.7%; Score 95.4; DB 9; Length 459;
Best Local Similarity 67.2%; Pred. No. 1.5e-15;
Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 126 ACCACTGATCTGTCGCTGCTTCCAGACCCCTTACCTGTAATCTGTGCTGGCTTCC 185
DB 406 ACCATGTTGCTGTGCTGCTTCCCTTCAACACCTCTTCAAGACCCCTGCTGCTTCA 347
QY 186 TGGCGCTGCGCTTACTCATCAAGCCCGAATCAGAAAGTGTGTGATGACCTGGAACCGG 245
DB 346 TAGGTTTACCTTACTCATGAAGCTTAGGAGACAGAAAGATGTTGGCGACCTGACCGGG 287
QY 246 CCGCGCTTGTGCTCCAAAGCGAGTGTACAACTCCGCGCGATGTGACGCTGG 305
DB 286 CCGAGGCTTAGCTTCCACCGCCAAAGTGTGAACATCTGGGCTTGTGAGCATCC 227
QY 306 TGGCGCCACTGCTGCTCTGG 326
DB 226 TCATGACCATTTCTGCTCATCG 206

```

```
RESULT 3
US-10-197-666A-67
; Sequence 67, Application US/10197666A
; Publication No. US20030092037A1
; GENERAL INFORMATION:
; APPLICANT: ASAH KASEI KABIUSIKI KAISYA
; TITLE OF INVENTION: E1k1 phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(389)
US-10-197-666A-67

Query Match
Best Local Similarity 62.8%; Pred. No. 5e-15; Length 621;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATGTCGTGGTGTTCACACCCCTTACCTGATGTCGTGGCTTC 184
DB 156 GACCACTGATGTCGTGGTGTTCACACCCCTTACCTGATGTCGTGGCTTC 215
QY 185 CTGGCCCTGGCCCTACCTCCATCAAGGCCGAGATCAAGAGTGTGTGACCTGGAAGC 244
DB 216 ATAGCACTTGGCCGACCTCCGAGGTCTAGGAGACAGAAAGTGTGGCGACGTGACCGG 275
QY 245 GCCCGCGCTTTTGGCTCCAAAGCCAGTGTACCAATCCTGGCCGCGATGTGAGAGC 304
DB 276 GCCCAGGCTATGCTCCACCCGCAAGTGTGAAACATCTGGGCCCTGATTTTGGGCATC 335
QY 305 GTGCCGCGCACTGTGCTCTGCTGGGGCTGTGTGACTGTGTGACCTGTGACCTG 355
DB 336 TTCATGACCATCTGTGCTCATCATCATCCAGTGTGTGTGTGTCAGGCCGAG 386

RESULT 4
US-09-969-708-88
; Sequence 88, Application US/09969708
; Patent No. US20020102532A1
; GENERAL INFORMATION:
; APPLICANT: Augeneus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-70
; CURRENT APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US/60/237,606
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,608
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,425
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 658
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 88
; LENGTH: 905
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-708-88

Query Match
Best Local Similarity 62.8%; Pred. No. 5.1e-15; Length 905;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATGTCGTGGTGTTCACACCCCTTACCTGATGTCGTGGCTTC 184
DB 442 GACCACTGATGTCGTGGTGTTCACACCCCTTACCTGATGTCGTGGCTTC 501
QY 185 CTGGCCCTGGCCCTACCTCCATCAAGGCCGAGATCAAGAGTGTGTGACCTGGAAGC 244
DB 502 ATAGCACTTGGCCGACCTCCGAGGTCTAGGAGACAGAAAGTGTGGCGACGTGACCGG 561
QY 245 GCCCGCGCTTTTGGCTCCAAAGCCAGTGTACCAATCCTGGCCGCGATGTGAGAGC 304
DB 562 GCCCAGGCTATGCTCCACCCGCAAGTGTGAAACATCTGGGCCCTGATTTTGGGCATC 621
QY 305 GTGCCGCGCACTGTGCTCTGCTGGGGCTGTGTGACTGTGTGACCTGTGACCTG 355
DB 622 TTCATGACCATCTGTGCTCATCATCATCCAGTGTGTGTGTGTCAGGCCGAG 672

RESULT 5
US-10-301-822-82
; Sequence 82, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kametkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen J.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MEM01-0292RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 905
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (280)...(678)
US-10-301-822-82

Query Match
Best Local Similarity 62.8%; Pred. No. 5.1e-15; Length 905;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATGTCGTGGTGTTCACACCCCTTACCTGATGTCGTGGCTTC 184
DB 442 GACCACTGATGTCGTGGTGTTCACACCCCTTACCTGATGTCGTGGCTTC 501
QY 185 CTGGCCCTGGCCCTACCTCCATCAAGGCCGAGATCAAGAGTGTGTGACCTGGAAGC 244
```


TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (280) ..(675)
US-10-197-666A-145

Query Match 14.3%; Score 93.4; DB 15; Length 905;
Best Local Similarity 62.8%; Pred. No. 5.1e-15;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATCTGTCGGTGTTCAGACCCCTTACCTGAATCTGTTGCTCGGCTTC 184
DB 442 GACCACTGTCGTGTCCTGTCCTGTTCAACCTCTTATGAACCTGCTGCTGGCTTC 501
QY 185 CTGGCGCTGGCTTCTCATCAAGGCCGAGATCAAGAGTGTGTGACCTGGAAGCG 244
DB 502 ATAGCATTCGCTACCTCCGTAAGTCTAAGGACAGAAAGTGTGTGCGACGTCGACCGG 561
QY 245 GCCCGCGCTTTGGCTTCCCAAGCCAGTGTACAAATCTGGCCGCGATGTGACGCTG 304
DB 562 GCCCAGGCTATGCTCCACCGCAGAGTGTGAAACATCTGGGCTCTGATTTTGGGCATC 621
QY 305 GTGCCGCACTGCTGCTCGGGGCTGTGTGACTGTGCTGACCTG 355
DB 622 TTCATGACCATTCCTGCTGCTCATCATCCAGTGTGTGTGTCGACGCCGAG 672

RESULT 9

US-10-197-666A-147
Sequence 147, Application US/1019766A
Publication No. US20030092037A1
GENERAL INFORMATION:
APPLICANT: ASAMI KASEI KABUSIKI KAISYA
TITLE OF INVENTION: E1K1 phosphorylation related gene
FILE REFERENCE: PH-1548US
CURRENT APPLICATION NUMBER: US/10/197,666A
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: JP 2001-218204
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: JP 2001-263450
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: JP 2002-012176
PRIOR FILING DATE: 2002-01-21
PRIOR APPLICATION NUMBER: US 60/305,884
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/316,304
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: US 60/350,027
PRIOR FILING DATE: 2002-01-23
NUMBER OF SEQ ID NOS: 156
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 147
LENGTH: 905
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (280) ..(675)
US-10-197-666A-147

Query Match 14.3%; Score 93.4; DB 15; Length 905;
Best Local Similarity 62.8%; Pred. No. 5.1e-15;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 125 GACCACTGATCTGTCGGTGTTCAGACCCCTTACCTGAATCTGTTGCTCGGCTTC 184
DB 442 GACCACTGTCGTGTCCTGTCCTGTTCAACCTCTTATGAACCTGCTGCTGGCTTC 501
QY 185 CTGGCGCTGGCTTCTCATCAAGGCCGAGATCAAGAGTGTGTGACCTGGAAGCG 244
DB 502 ATAGCATTCGCTACCTCCGTAAGTCTAAGGACAGAAAGTGTGTGCGACGTCGACCGG 561

QY 245 GCCCGCGCTTTGGCTTCCAAAGCCAAAGTCTACAAATCTTGGCCGCGATGTGACGCTG 304
DB 562 GCCCAGGCTATGCTCCACCGCAGAGTGTCTGAACATCTGGGCTCTGATTTTGGGCATC 621
QY 305 GTGCCGCACTGCTGCTCGGGGCTGTGTGACTGTGACCTGTGACCTG 355
DB 622 TTCATGACCATTCCTGCTCATCATCCAGTGTGTGTGTGTCGACGCCGAG 672

RESULT 10

US-10-115-831-27
Sequence 27, Application US/10115831
Publication No. US20030219743A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyao
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20030219743A1el Nucleic Acids and
FILE REFERENCE: Polypeptides
FILE REFERENCE: 792CIP2ADIV
CURRENT APPLICATION NUMBER: US/10/115,831
CURRENT FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: 09/667,298
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 178
SOFTWARE: pt_FL_genes Version 2.0
SEQ ID NO 27
LENGTH: 920
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (280) ..(678)
US-10-115-831-27

Query Match 14.3%; Score 93.4; DB 13; Length 920;
Best Local Similarity 62.8%; Pred. No. 5.1e-15;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTGATCTGTCGGTGTTCAGACCCCTTACCTGAATCTGTTGCTCGGCTTC 184
DB 442 GACCACTGTCGTGTCCTGTCCTGTTCAACCTCTTATGAACCTGCTGCTGGCTTC 501
QY 185 CTGGCGCTGGCTTCTCATCAAGGCCGAGATCAAGAGTGTGTGACCTGGAAGCG 244
DB 502 ATAGCATTCGCTACCTCCGTAAGTCTAAGGACAGAAAGTGTGTGCGACGTCGACCGG 561
QY 245 GCCCGCGCTTTGGCTTCCAAAGCCAAAGTCTACAAATCTTGGCCGCGATGTGACGCTG 304
DB 562 GCCCAGGCTATGCTCCACCGCAGAGTGTCTGAACATCTGGGCTCTGATTTTGGGCATC 621
QY 305 GTGCCGCACTGCTGCTCGGGGCTGTGTGACTGTGACCTGTGACCTG 355
DB 622 TTCATGACCATTCCTGCTCATCATCCAGTGTGTGTGTGTCGACGCCGAG 672

RESULT 11

US-10-198-846-10182
Sequence 10182, Application US/10198846
Publication No. US20030099974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinhilber, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049

CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 10182
LENGTH: 1707
TYPE: DNA
ORGANISM: Homo sapiens
US-10-198-846-10182

Query Match 14.3%; Score 93.4; DB 15; Length 1707;
Best Local Similarity 62.8%; Pred. No. 5.2e-15;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy 125 GACCACTGATCTGTGCTGTTGAGCACCCTTACCTGATCTGTGTTGCTGGCTTC 184
Db 441 GACCACTGCTGTGCTGCTGCTTCAACACCTCTTCAATGAACTGCTGCTGCTTC 500
Qy 185 CTGGCGCTGCTCTACCTCATCAAGCCCGAGATCAAGAGTGTGTGTAACCTGAACG 244
Db 501 ATAGCATTCGCTACTCCGTAAAGTCTAGGACAGAAAGATGTTGGGACGTACCGGG 560
Qy 245 GCGCGGCTTTGGTCCAAAGCCAAAGTCTACCAACATCTGCGCCGATGTGAAGCTG 304
Db 561 GCGCGGCTTTGGTCCAAAGCCAAAGTCTGGAACATCTGCGCCGATTTTGGGATC 620
Qy 305 GTGCGGCACTGCTGCTCTGAGGCTGTGTGTAAGTGTGCTGCACTG 355
Db 621 TTCAATGACATTCCTGCTGTCATCATCCCAATGTTGTCGTCAGGCCAG 671

RESULT 12
US-09-918-995-6813
Sequence 6813, Application US/09918995
Publication No. US2003007623A1
GENERAL INFORMATION:
APPLICANT: HySeq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 6813
LENGTH: 413
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(413)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-6813

Query Match 14.3%; Score 93.2; DB 11; Length 413;
Best Local Similarity 66.3%; Pred. No. 5.5e-15;
Matches 134; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 125 GACCACTGATCTGTGCTGTTGAGCACCCTTACCTGATCTGTGTTGCTGGCTTC 184
Db 187 GACCACTGCTGTGCTGCTGCTTCAACACCTCTTCAATGAACTGCTGCTGCTTC 246
Qy 185 CTGGCGCTGCTCTACCTCATCAAGCCCGAGATCAAGAGTGTGTGTAACCTGAACG 244
Db 247 ATAGCATTCGCTACTCCGTAAAGTCTAGGACAGAAAGATGTTGGGACGTACCGGG 306
Qy 245 GCGCGGCTTTGGTCCAAAGCCAAAGTCTACCAACATCTGCGCCGATGTGAAGCTG 304
Db 307 GCGCGGCTTTGGTCCAAAGCCAAAGTCTGGAACATCTGCGCCGATTTTGGGATC 366

Qy 305 GTGCGGCACTGCTGCTCTG 326
Db 367 CTATGACATTCCTGCTCATCG 388

RESULT 13
US-10-115-831-9
Sequence 9, Application US/10115831
Publication No. US20030219743A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20030219743A1 Nucleic Acids and
FILE REFERENCE: 792CIP2ADIV

CURRENT APPLICATION NUMBER: US/10/115,831
CURRENT FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: 09/667,298
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 178
SOFTWARE: pt_gene Version 2.0
SEQ ID NO: 9
LENGTH: 596
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (22)-(423)
US-10-115-831-9

Query Match 14.3%; Score 93.2; DB 13; Length 596;
Best Local Similarity 66.3%; Pred. No. 5.6e-15;
Matches 134; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 125 GACCACTGATCTGTGCTGTTGAGCACCCTTACCTGATCTGTGTTGCTGGCTTC 184
Db 187 GACCACTGCTGTGCTGCTGCTTCAACACCTCTTCAATGAACTGCTGCTGCTTC 246
Qy 185 CTGGCGCTGCTCTACCTCATCAAGCCCGAGATCAAGAGTGTGTGTAACCTGAACG 244
Db 247 ATAGCATTCGCTACTCCGTAAAGTCTAGGACAGAAAGATGTTGGGACGTACCGGG 306
Qy 245 GCGCGGCTTTGGTCCAAAGCCAAAGTCTACCAACATCTGCGCCGATGTGAAGCTG 304
Db 307 GCGCGGCTTTGGTCCAAAGCCAAAGTCTGGAACATCTGCGCCGATTTTGGGATC 366
Qy 305 GTGCGGCACTGCTGCTCTGAGGCTGTGTGTAAGTGTGCTGCACTG 355
Db 367 CTATGACATTCCTGCTCATCG 388

RESULT 14
US-10-197-666A-71
Sequence 71, Application US/10197666A
Publication No. US20030092037A1
GENERAL INFORMATION:
APPLICANT: ASAH KASEI KANUSIKI KAISYA
TITLE OF INVENTION: B1KI phosphorylation related gene
FILE REFERENCE: PH-1548US
CURRENT APPLICATION NUMBER: US/10/197,666A
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: JP 2001-218204
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: JP 2001-263450
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: JP 2002-012176
PRIOR FILING DATE: 2002-01-21

Qy 125 GACCACTGATCTGTGCTGTTGAGCACCCTTACCTGATCTGTGTTGCTGGCTTC 184
Db 187 GACCACTGCTGTGCTGCTGCTTCAACACCTCTTCAATGAACTGCTGCTGCTTC 246
Qy 185 CTGGCGCTGCTCTACCTCATCAAGCCCGAGATCAAGAGTGTGTGTAACCTGAACG 244
Db 247 ATAGCATTCGCTACTCCGTAAAGTCTAGGACAGAAAGATGTTGGGACGTACCGGG 306
Qy 245 GCGCGGCTTTGGTCCAAAGCCAAAGTCTACCAACATCTGCGCCGATGTGAAGCTG 304
Db 307 GCGCGGCTTTGGTCCAAAGCCAAAGTCTGGAACATCTGCGCCGATTTTGGGATC 366

